You Can't Always Get What You Want: Tobacco Excises and the "Magic Bullet" in Bolivia [*]

Rómulo A. Chumacero [†]

Abstract

Tobacco excises are often expected to reduce smoking, raise revenue, and improve welfare. This paper argues that this claim fails as a general proposition when legal and untaxed cigarettes coexist and consumers can substitute across brands and channels. Using monthly brand-level data for Bolivia's legal cigarette market, the paper studies the 2018 reform that replaced an *ad valorem* excise with category-specific *ad quantum* taxation. The reform changed relative effective tax burdens: lower-price legal brands faced higher effective tax shares, while some higher-price brands faced lower burdens. Gross real prices show no statistically significant post-reform increase; adjustment appears mainly in net-of-tax prices, with margin compression in lower-price brands and positive net-price adjustments in higher-price brands. Brand-level legal demand is much more price-responsive than aggregate legal demand, showing that aggregation conceals substitution margins. In the observed legal market, legal sales and real excise revenue both fell, while formal-chain indicators weakened. The results reject the fiscal component of the strong "magic bullet" claim in the observed legal segment but do not identify total cigarette consumption. A welfare break-even exercise shows that, even under assumptions favorable to the reform, the required uninternalized harm is about 30 percent of the average legal price per cigarette and rises quickly if legal-base erosion reflects substitution toward untaxed products. Excise reforms must therefore be evaluated using the taxed base that remains after consumers and firms adjust, not statutory tax rates alone.

Keywords: Tobacco excises; taxed legal base; illicit substitution; cigarette demand; price adjustment; Bolivia.

JEL Classification: C23, D12, H21, I18, L66.

First version: April 2026

This version: May 2026

---

[*] I would like to thank Ricardo Paredes for useful comments and suggestions. I gratefully acknowledge financial support from CITSA and access to the data used in this study. CITSA had no role in the analysis, interpretation of the results, or writing of the manuscript. The views expressed in this paper, as well as any remaining errors, are solely my responsibility.

[†] Department of Economics. University of Chile. E-mail: rchumace@fen.uchile.cl.

1 Introduction

Tobacco excise taxation is often presented as a policy with unusually attractive properties. By raising cigarette prices, it should reduce smoking; by increasing the tax collected on each unit sold, it should also raise fiscal revenue. In that sense, tobacco excises are frequently treated as a "magic bullet": a single instrument that may improve public health while strengthening the fiscal position of the state.

That argument is appealing, but it is not mechanical. A statutory tax change is only the starting point of a sequence of market adjustments. Firms may pass the tax through to final prices fully, partially, or more than fully, depending on costs, market power, and competitive pressure. Consumers may reduce legal purchases, but they may also substitute across legal brands, trade down to cheaper products, or shift purchases toward untaxed and illicit channels. The tax authority collects revenue only on legal tax-paid sales, not on total cigarette consumption. Thus, the fiscal success of an excise increase depends not simply on whether the tax per unit rises, but on whether the legal tax base contracts less than the increase in the effective tax burden.

This distinction is the central object of the paper. Public-health arguments are naturally framed in terms of total cigarette consumption. Fiscal arguments, by contrast, depend on the taxed legal base. These two objects coincide only under restrictive conditions: legal sales must track total consumption closely, consumers must not substantially substitute toward untaxed alternatives, and the effective tax burden must translate into revenue rather than into erosion of the base. Once those conditions fail, a decline in legal sales has an ambiguous implication for total smoking: total consumption may fall by less than legal sales, may remain unchanged, or may even increase if substitution toward untaxed products is sufficiently strong.

Bolivia provides a useful setting for studying these mechanisms. In 2018, the cigarette excise shifted from an *ad valorem* structure to a category-specific *ad quantum* regime. Under the new system, tax liability depended on physical units rather than value. As a result, within each tax category, the effective tax share was mechanically higher for lower-

price legal brands. This matters because lower-price legal brands are part of the segment in which substitution across legal brands and toward untaxed alternatives may be especially relevant for revenue. In a differentiated cigarette market, the design of the tax may therefore affect relative prices, brand substitution, the composition of legal sales, and the vulnerability of the taxed base.

This paper evaluates the "magic bullet" claim for Bolivia using monthly brand-level data on legal cigarette sales, prices, excise payments, and related market variables for the period 2015–2025. The empirical analysis begins from a simple fact: in the observed legal market, the strong version of the claim did not materialize. Legal cigarette sales declined sharply, but real excise revenue also fell. This is enough to reject the fiscal component of the strong claim for the observed legal base. The more difficult question is how to interpret the decline in legal sales. If legal purchases fell because total smoking fell by the same amount, then the public-health interpretation would be straightforward. If instead part of the adjustment occurred through substitution across legal brands or toward untaxed alternatives, the contraction of the legal market would overstate the reduction in smoking. If substitution toward untaxed products was sufficiently strong, total consumption may have fallen less than legal sales, remained unchanged, or even increased.

The main empirical result is that aggregation hides precisely the margins that matter for this question. When demand is estimated at the aggregate level, legal cigarette demand appears relatively inelastic. When the same problem is studied at the brand level, demand is much more price-responsive. This contrast is not a minor econometric detail. It indicates that consumers adjust along margins that aggregate data are poorly equipped to identify. It also has implications for studies based on household expenditure surveys, where prices are often measured as unit values. If unit values combine true prices with brand choice, quality choice, reporting error, and legal-illicit substitution, the resulting elasticities may be poor guides to the response of the taxed legal base.

The paper also studies demand dynamics. A lagged-demand specification produces results broadly consistent with the static evidence and allows short-run and long-run elasticities to be compared. A forward-looking rational-addiction specification is also estimated, but

the available instruments do not identify that model with enough precision to make its implied elasticities useful for policy evaluation. The dynamic analysis therefore reinforces the main conclusion rather than replacing it: strong policy claims should not rest on weakly identified dynamic specifications.

Taken together, the evidence is difficult to reconcile with the strong "magic bullet" view. The observed legal market contracted sharply, but that contraction was not associated with higher real excise revenue. Brand-level demand is much more price-responsive than aggregate demand, while the evolution of prices, legal sales, and formal-sector indicators shows that the contraction of the legal market extended beyond tax payments. These patterns are more consistent with a setting in which the legal tax base is vulnerable to substitution across brands and, plausibly, toward untaxed alternatives.

The paper does not claim to observe total cigarette consumption directly, nor does it identify the full size of the illicit market. Its more precise claim is that legal sales cannot be treated as total consumption, that aggregate demand elasticities are not sufficient for fiscal evaluation, and that welfare conclusions require additional assumptions about externalities, internalities if admitted, enforcement costs, and the social value of public funds.

The paper makes four contributions. First, it distinguishes total cigarette consumption from the taxed legal base and shows why the latter is the relevant object for fiscal evaluation. Second, it documents that aggregate cigarette elasticities can conceal substantial substitution across brands, making them unreliable for evaluating excise revenue. Third, it evaluates demand dynamics while distinguishing reduced-form persistence from forward-looking rational addiction. Fourth, it links demand, tax design, legal-base erosion, and formal-sector outcomes in a unified framework that clarifies the limits of welfare conclusions based only on legal-market data.

The remainder of the paper is organized as follows. Section 2 reviews the literature by mechanism. Section 3 develops the theoretical framework. Section 4 describes the Bolivian

institutional setting, the tax reform, and the data. Section 5 presents the empirical analysis. Section 6 concludes.

2 Related Literature and Mechanisms

Tobacco excises are usually studied through separate empirical margins: price measurement, demand elasticities, illicit trade, pass-through, fiscal revenue, production effects, and welfare. Each margin is important, but the "magic bullet" claim is a joint claim. It requires that an excise increase be transmitted to consumer prices, reduce total tobacco consumption, preserve enough of the taxed legal base to raise revenue, limit substitution toward untaxed channels, and improve welfare once private utility, externalities, internalities, enforcement, and public spending are considered. This section reviews the literature by mechanism. The purpose is not to provide an exhaustive survey, but to identify why each margin matters and why evaluating tobacco excises requires considering them jointly.

2.1 Expenditure surveys and the measurement of prices

A common approach to estimating cigarette demand, especially in countries where scanner data or administrative transaction data are limited, is to rely on household expenditure surveys such as Household Income and Expenditure Surveys (HIES) and Living Standards Measurement Study (LSMS) surveys. These surveys are attractive because they contain household-level information on expenditures, quantities, income, demographics, and location. Their main limitation for tobacco-tax analysis is that they typically do not report transaction prices. Analysts therefore construct unit values by dividing expenditure by reported quantity.

The difficulty is that unit values are not prices. They combine true price variation with quality choice, brand composition, pack-size differences, reporting error, and possibly the use of legal and illegal purchasing channels. This issue is central in Deaton's analysis of quality, quantity, and spatial price variation (Deaton, 1988). The Deaton approach treats

unit values as endogenous objects and exploits spatial variation, usually at the cluster level, to separate price variation from household quality choice and measurement error. In tobacco applications, John et al. (2023) provide a widely used adaptation of this logic for household expenditure surveys. Related applications include Morales and Santander (2025) for Bolivia, Paraje et al. (2021) for El Salvador, and Dare et al. (2021) for South Africa.

This strategy is useful, but it relies on strong identifying assumptions. A key requirement is that most of the relevant price variation occurs across clusters, while within-cluster variation in unit values reflects quality choice and measurement error, rather than true price differences or channel choice. This assumption becomes demanding in differentiated cigarette markets. Consumers can switch across brands, quality segments, pack formats, legal outlets, informal sellers, and illicit products. If a household responds to higher legal prices by moving to a cheaper brand or channel, the unit value already reflects behavioral adjustment. Treating it as an exogenous price may therefore confound the cause and the consequence of substitution.

The implication for tobacco excise evaluation is broader than a standard measurement-error concern. Gibson and Kim (2019) show that quality responses can materially affect estimated price elasticities. For fiscal policy, the revenue-relevant object is the response of tax-paid legal sales to the legal tax-inclusive price. A unit value constructed from household expenditure may instead reflect a mixture of legal purchases, illicit purchases, quality downgrading, and brand switching. Even when the resulting elasticity is internally coherent for reported household purchases, it need not measure the elasticity of the taxed legal base. This matters for the fiscal claim because fiscal revenue depends on the taxed legal base, while household unit values may partly absorb the substitution margins that erode that base.

2.2 Demand elasticities, dynamics, and addiction

The empirical literature on cigarette demand reports a wide range of price elasticity estimates. Surveys such as Badenes-Plá and Jones (2003), Chaloupka and Warner (2000), and DeCicca et al. (2022) show that estimates differ across countries, income groups, age

groups, data sources, and econometric designs. This dispersion should not be interpreted only as disagreement. It often reflects differences in the object being measured. Some studies estimate participation responses, others conditional intensity. Some use aggregate sales, others household purchases. Some measure short-run responses, others long-run effects. Some observe legal tax-paid sales, while others observe reported purchases that may combine legal and illicit sources.

These distinctions matter for excise evaluation because the joint claim combines several different demand objects. Public-health analysis requires evidence on total tobacco consumption and its health consequences. Fiscal analysis requires evidence on the tax-paid legal base. These objects coincide only when legal sales track total consumption closely. Firm pricing depends on residual demand for each brand or market segment. Welfare analysis depends on how consumption responses interact with externalities, internalities, fiscal revenue, enforcement, and substitution across legal and untaxed alternatives. An elasticity estimated for one object cannot be automatically used to evaluate another.

Demand dynamics add a further distinction. Cigarette consumption may be persistent because of habits, addiction, adjustment costs, or unobserved heterogeneity. Becker et al. (1994) provide the canonical rational-addiction framework, in which current consumption is linked to past consumption, expected future consumption, and expected future prices. Orphanides and Zervos (1995) extend the analysis by allowing learning and regret. These models imply that short-run and long-run responses may differ. At the same time, lagged consumption in an empirical demand equation should not be interpreted mechanically as evidence of addiction. As Gilleskie (2014) emphasizes, dynamic demand models must distinguish true state dependence from persistent unobserved heterogeneity.

Recent empirical evidence reinforces the need to distinguish horizons, data frequency, and market segments. Ferre et al. (2025), using monthly national data for Uruguay, report differences between short-run and long-run cigarette-demand responses, consistent with gradual adjustment. Tauras et al. (2016) show that price responsiveness may vary across the price distribution, which cautions against extrapolating from average elasticities when excise reforms change relative prices across products.

The implication is that dynamic structures may matter, but identification matters as much as theory. Static demand estimates can provide useful conditional benchmarks. Lagged-demand specifications can summarize persistence and allow short-run and long-run elasticities to differ. Forward-looking rational-addiction models are conceptually richer, but empirically demanding because future consumption and future prices must be instrumented with variables in the consumer's information set. For this reason, the empirical analysis below distinguishes these specifications rather than treating all price elasticities as interchangeable.

2.3 Total consumption, tax-paid sales, and illicit substitution

A central difficulty in tobacco-tax evaluation is the distinction between total cigarette consumption and tax-paid legal sales. The former is the natural object for public-health analysis. The latter is the object that determines excise revenue. These objects coincide only under restrictive conditions: all consumption must be legal and taxed, or the legal share of total consumption must remain sufficiently stable. Once consumers can shift across products or purchasing channels, legal sales become a mixture of two margins: changes in total smoking and changes in the legal share of consumption.

This distinction is familiar from the broader public-finance literature. The taxable base is not a mechanical object; it responds to statutory rates, enforcement, administrative capacity, and avoidance technologies. Slemrod and Yitzhaki (2002) emphasize that avoidance, evasion, and administration are central to tax analysis because the relevant behavioral response is often the response of the tax base, not only the response of the underlying activity. In cigarette markets, this distinction is especially important because the same statutory tax change may affect legal tax-paid purchases, total smoking, and illicit-market activity in different ways.

Several studies document margins through which cigarette taxes can separate legal sales from total consumption. DeCicca et al. (2013) study legal excise-tax avoidance through price search and lower-tax purchasing opportunities, while Lovenheim (2008) shows that

cross-border shopping can materially affect cigarette-tax revenues. Gruber et al. (2003) show that inference about cigarette price responsiveness can be sensitive to the treatment of smuggling. Jha and Chaloupka (2000) discuss the role of illicit trade and enforcement in tobacco-control policy, particularly in developing-country contexts. Joossens and Raw (2008) emphasize the supply-chain and enforcement dimensions of cigarette smuggling, and Saenz de Miera Juarez et al. (2021) show that measuring illicit cigarette markets requires triangulation across methods.

For the fiscal claim, this distinction is decisive. Revenue is collected on legal tax-paid sales, not on total cigarette consumption. Health outcomes depend on total consumption, not on the legal base alone. Therefore, an excise increase can fail fiscally if the legal base erodes, even if total consumption falls. Conversely, legal sales may decline while total consumption falls less, remains roughly unchanged, or even increases if substitution toward untaxed products is sufficiently strong. This is not a claim that illicit substitution necessarily dominates. It is a reminder that the sign and magnitude of the total-consumption response cannot be inferred from legal sales alone when avoidance, evasion, or illicit substitution are relevant margins.

2.4 Pass-through, market structure, and excise design

Excise taxes affect consumers through the prices and opportunities they face. The statutory tax is therefore not equivalent to the effective consumer-price change. The mapping from taxes to retail prices depends on pass-through, costs, market structure, product differentiation, and competitive pressure from substitute products, including untaxed alternatives. This matters for both sides of the "magic bullet" claim: the health margin depends on the consumer-price response, while the fiscal margin depends on the tax payments collected on the legal base that remains after consumers and firms adjust.

Pass-through is an equilibrium outcome rather than a mechanical accounting identity. Under imperfect competition, firms may absorb part of a tax increase, shift it fully to consumers, or raise final prices by more than the tax increase. Delipalla and O'Donnell (2001) show that tax incidence, market power, and conduct affect cigarette pricing in the

European cigarette industry. The implication is that statutory tax changes cannot be interpreted directly as changes in consumer prices. If pass-through is incomplete, the price response faced by consumers may be smaller than the statutory change suggests. If pass-through exceeds one, consumer prices may rise more than proportionally, but the fiscal effect still depends on the response of tax-paid legal sales.

The form of the excise also matters. *Ad valorem* taxes are proportional to value, while *ad quantum* taxes are tied to physical units. In differentiated markets, this distinction can affect relative prices and the composition of sales. A unit tax represents a larger effective tax share for lower-price products within a tax category, while an *ad valorem* tax preserves proportionality with value. Thus, a shift from value-based to unit-based taxation can change the relative burden across products even if the average statutory burden changes only modestly. When consumers can substitute across brands, quality segments, or legal and untaxed channels, these relative-price effects may be central for the response of the taxed legal base.

Evidence from related nicotine-product markets reinforces the broader point that tax changes can alter both prices and product choice. Cotti et al. (2022), using retail panel data on e-cigarette taxes, document pass-through and cross-product substitution between e-cigarettes and cigarettes. The relevance for tobacco excise evaluation is not that those estimates carry over directly to Bolivia, but that price transmission and substitution must be studied jointly. A tax reform first changes the effective price vector; firms and consumers then adjust along product and channel margins; fiscal revenue depends on the legal base that remains.

This mechanism is important for the empirical analysis below. The paper does not estimate a full structural model of pass-through. Instead, it examines the evolution of effective tax burdens, gross prices, net-of-tax prices, and legal quantities across brands. That evidence is useful because excise policy cannot be evaluated from statutory tax rates alone: the relevant margins are the prices consumers face, the net-of-tax prices associated with legal sales, and the legal quantities on which revenue is collected.

2.5 Formal production, employment, and the legal chain

Tax-paid cigarette sales are embedded in a formal production and distribution chain. A contraction in legal sales may therefore affect not only final consumption and tax payments, but also production, employment, labor costs, distribution activity, and purchases from domestic input suppliers. These margins matter for the "magic bullet" claim because the fiscal promise of excise taxation depends on the resilience of the legal tax base while the market adjusts.

The literature is appropriately cautious in interpreting these effects. Sectoral contractions do not translate mechanically into aggregate welfare losses: labor, expenditure, and fiscal resources may be reallocated across sectors. Cicowiez et al. (2024), using a computable general-equilibrium model for Argentina, show that the employment effects of higher tobacco taxes depend on reallocation mechanisms and fiscal-closure assumptions. This perspective is useful because it prevents partial-equilibrium declines in tobacco-related activity from being interpreted as aggregate employment or welfare losses.

At the same time, the formal legal chain remains relevant for evaluating excise policy. If higher taxes are accompanied by lower tax-paid sales, lower real revenue, and weaker formal activity, the adjustment is broader than a change in consumer prices or smoking behavior. It affects the part of the market through which taxes are collected and legal activity is organized. This is especially important when legal products compete with untaxed alternatives: erosion of the legal base may reduce revenue, weaken formal activity, and shift transactions outside the observable and taxable market. Formal-sector indicators therefore help assess whether the contraction of legal sales reflects a narrow demand response or a broader weakening of the legal taxed segment.

2.6 Welfare, incidence, and corrective taxation

The "magic bullet" claim is ultimately a welfare claim, not only a prediction about prices, quantities, or revenue. Under the standard corrective-tax rationale, tobacco taxation may

be justified by externalities, such as secondhand smoke, publicly financed health costs, or other costs imposed on non-smokers and taxpayers. In that case, the welfare case for taxation depends on whether the tax reduces the harmful activity and whether the external costs avoided exceed the private and fiscal costs generated by the policy.

Behavioral public finance adds a second possible rationale. If smokers do not fully internalize future harms because of addiction, self-control problems, or time inconsistency, taxation may also address internalities. Gruber and Kőszegi (2004) show that the welfare incidence of cigarette taxes changes when time inconsistency and self-control benefits are incorporated. O'Donoghue and Rabin (2006) formalize optimal sin taxes with self-control problems, while Farhi and Gabaix (2020) provide a broader framework for optimal taxation with behavioral agents. Jürges (2004) further highlights that the welfare costs of addiction may extend beyond narrowly measured health consequences. These arguments make explicit that welfare conclusions depend on normative and behavioral assumptions, not only on observed legal sales.

Sufficient-statistics approaches clarify why the relevant behavioral margin matters. Chetty (2009) emphasizes that elasticities summarize welfare-relevant responses only under clear assumptions about the tax base and the behavior being measured. In tobacco markets, this distinction is crucial. The elasticity of total cigarette consumption is relevant for externalities and internalities; the elasticity of tax-paid legal sales is relevant for revenue; and the elasticity of the legal base may partly reflect avoidance or illicit substitution rather than lower smoking. DeCicca et al. (2013) show that avoidance margins affect optimal corrective taxation precisely because the taxable base can respond differently from the harmful activity.

Illicit substitution further complicates welfare evaluation. Becker and Murphy (2006) emphasize that illegal markets introduce enforcement considerations, resource costs, and rents. If higher taxes move consumers from legal to untaxed cigarettes, the policy may reduce revenue, raise enforcement or resource costs, weaken formal activity, and shift transactions outside the observable and taxable market. This does not imply that the tax reduces welfare. It implies that a welfare gain cannot be inferred from lower legal sales or

lower total consumption alone. The relevant question is whether the externality or internality gains associated with any reduction in total smoking are large enough to offset lost private surplus, lower fiscal revenue, enforcement or resource costs, and any costs associated with displacement toward untaxed products. The paper therefore treats welfare as an accounting and interpretation problem: the empirical results identify important legal-market margins, but a full welfare ranking requires assumptions about objects that are not directly observed.

3 A unified framework for excise evaluation

The previous section reviewed the main mechanisms through which tobacco excises operate: price measurement, demand responses, legal-base erosion, price adjustment, formal-sector effects, and welfare. This section brings those mechanisms into a single framework. The purpose is not to estimate a full structural model of the cigarette market, but to discipline the objects that must be distinguished when evaluating an excise reform. The framework separates four margins: the effective consumer prices generated by taxes and firm pricing, total cigarette consumption, tax-paid legal sales, and welfare. This distinction is central because the "magic bullet" claim requires more than a statutory tax increase. It requires tracing how taxes affect relevant prices, how consumers reallocate across legal and untaxed alternatives, how the taxed legal base responds, and what additional assumptions are needed to draw welfare conclusions.

3.1 Consumer demand and dynamic adjustment

Consider an infinitely lived consumer who, in each period t, chooses legal cigarette consumption by brand, illicit cigarette consumption, other goods, numeraire consumption, and savings. Let $q_{n,t}$ denote consumption of legal cigarette brand n=1,..,N, and let q_t^I denote consumption of the illicit alternative. Legal consumer prices are denoted by $p_{n,t}$, and the effective illicit price is denoted by p_t^I. The latter should be interpreted broadly: it includes not only the money price of illicit cigarettes, but also search costs, availability, quality uncertainty, risk, and enforcement-related costs. Let o_t denote a vector of other

goods, with price vector w_t, and let m_t be the numeraire. Income, expressed in units of the numeraire, is denoted by y_t. The variable a_{t+1} denotes savings carried to period $t+1$ and earning net return r_{t+1} in that period; s_t is the stock of the potentially addictive consumption of cigarettes, depreciating at rate $\delta \in (0,1]$; $\beta \in (0,1)$ is the subjective discount factor; and $\mathbb{E}_t$ denotes expectations conditional on information available at date t.

The consumer maximizes:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u\left(q_{1,t},..,q_{N,t},q_t^I,o_t,m_t,s_t\right), \tag{1}$$

subject to the budget constraint:

$$y_t + \left(1+r_t\right)a_t = \sum_{n=1}^{N} p_{n,t}q_{n,t} + p_t^I q_t^I + w_t{}' o_t + m_t + a_{t+1}, \tag{2}$$

and the habit-stock law of motion:

$$s_{t+1} = \left(1-\delta\right)s_t + \sum_{n=1}^{N} q_{n,t} + q_t^I. \tag{3}$$

The stock s_t summarizes the dynamic consequences of past cigarette consumption. It may capture reinforcement, withdrawal, tolerance, or other forms of state dependence. The formulation is deliberately general. It provides a dynamic benchmark for the empirical analysis, but it does not require that all empirical specifications be interpreted as fully structural rational-addiction models.

Let λ_t be the multiplier on the budget constraint (2) and μ_t the shadow value associated with the habit-stock equation (3). Since m_t is the numeraire, the first-order condition for m_t implies $u_{m_t} = \lambda_t$. For legal brand n, the first-order condition is:

$$\mu_t = u_{q_{n,t}} - u_{m_t} p_{n,t}. \tag{4}$$

For the illicit alternative, the analogous condition is:

$$\mu_t = u_{q_t^I} - u_{m_t} p_t^I. \tag{5}$$

Equations (4) and (5) show that legal and illicit cigarette choices are linked by the common shadow value of the habit stock. Relative prices, marginal utilities, and the effective cost of illicit access jointly determine substitution across legal brands and between legal and illicit cigarettes.

The Euler equation associated with the habit stock is:

$$\mu_t = \beta\left(1-\delta\right)\mathbb{E}_t\mu_{t+1} - \beta\mathbb{E}_t u_{s_{t+1}}. \tag{6}$$

Combining (4) and (6) yields the forward-looking condition for legal brand n:[1]

$$u_{q_{n,t}} - u_{m_t} p_{n,t} = \beta\left(1-\delta\right)\mathbb{E}_t\left[u_{q_{n,t+1}} - u_{m_{t+1}} p_{n,t+1}\right] - \beta\mathbb{E}_t u_{s_{t+1}}. \tag{7}$$

Equation (7) is the dynamic core of the consumer problem. It implies that current legal cigarette demand may depend on current prices, expected future prices, and expected future consumption through the marginal-utility terms. A natural forward-looking empirical counterpart therefore includes current and future quantities and prices. Lagged consumption is not a direct regressor in (7); it enters through the habit stock s_t. Empirically, one may approximate this state variable with lagged consumption, but that corresponds to a reduced-form lagged-demand specification rather than to the most direct empirical counterpart of the Euler equation.

This distinction guides the empirical work. A static demand equation provides a conditional benchmark. A lagged-demand specification captures persistence and allows short-run and long-run elasticities to differ. A forward-looking rational-addiction specification follows the logic of (7) by including future consumption and future prices, which must be treated as endogenous and instrumented with variables in the information set available at time t. In all cases, the observed quantities are legal tax-paid sales by brand, so the estimated demand equations describe the legal segment of the market rather than total cigarette consumption.

[1] An analogous forward-looking condition for the illicit alternative, parallel to (7), is obtained by combining (5) and (6).

3.2 Legal supply, excise design, and price transmission

Consumer prices are not mechanical transformations of statutory taxes. They are equilibrium objects determined by taxes, costs, market structure, and substitution possibilities. This is especially relevant in differentiated cigarette markets, where the effect of a tax reform depends not only on the average tax burden but also on how the reform changes relative prices across brands and product categories.

Let each legal brand n be produced with input vector $x_{n,t}$ under a technology $q_{n,t} = F\left(x_{n,t}\right)$. Let v_t denote the vector of input prices. Cost minimization yields a conditional cost function $C_n\left(q_{n,t}, v_t\right)$, input demands $x_{n,t} = X_n\left(q_{n,t}, v_t\right)$, and marginal cost

$$c_{n,t} = \frac{\partial C_n\left(q_{n,t}, v_t\right)}{\partial q_{n,t}}.$$

Legal cigarette taxation may include both *ad valorem* and *ad quantum* components. Let $\gamma_{n,t}$ denote the *ad valorem* rate and $\tau_{n,t}$ the *ad quantum* tax applicable to brand n. Allowing these terms to vary by brand or category is useful because some tax systems apply different *ad quantum* amounts to different cigarette types. If $b_{n,t}$ denotes the net-of-tax producer price, the consumer price satisfies:

$$p_{n,t} = \left(1 + \gamma_{n,t}\right) b_{n,t} + \tau_{n,t}. \tag{8}$$

This formulation accommodates both *ad valorem* and *ad quantum* regimes. Under an *ad valorem* tax, the tax burden is proportional to value. Under an *ad quantum* tax, the statutory liability is tied to physical units. Therefore, within a given tax category, a unit tax represents a larger effective tax share for lower-price products. A shift from *ad valorem* to *ad quantum* taxation can therefore change relative tax burdens and relative prices across brands even when the average statutory burden changes modestly.

Consider a firm f that controls a set of brands $\mathcal{M}_f$. Demand for each brand depends on the full vector of legal prices, the effective illicit price, and demand shifters z:

$$q_{n,t} = D_n\left(p_{1,t},..,p_{N,t},p_t^I,z_t\right).$$

The firm's period-t profit is:

$$\pi_{f,t} = \sum_{n\in\mathfrak{N}_f}\left[b_{n,t}q_{n,t} - C_n\left(q_{n,t},v_t\right)\right].$$

The firm chooses consumer prices, taking the prices outside of $\mathfrak{N}_f$ as given. Differentiating profits with respect to the prices of the brands controlled by the firm gives the pricing condition:

$$A_f q_{f,t} + \Omega_f'\left(b_{f,t} - c_{f,t}\right) = 0,$$

where $q_{f,t}$, $b_{f,t}$, $c_{f,t}$ are the vectors of quantities, net-of-tax producer prices, and marginal costs for the firm's brands; A_f is a diagonal matrix with entries $1/\left(1+\gamma_{n,t}\right)$; and Ω_f is the Jacobian matrix of demand derivatives for the brands controlled by the firm, with generic element $\left(\partial q_{j,t} / \partial p_{n,t}\right)$.

Solving the pricing condition and using (8), consumer prices can be written as:

$$p_{f,t} = \tau_{f,t} + \Gamma_f\left[c_{f,t} - \left(\Omega_f'\right)^{-1} A_f q_{f,t}\right], \tag{9}$$

where Γ_f is a diagonal matrix with entries $1+\gamma_{n,t}$, and $\tau_{f,t}$ is the vector of *ad quantum* taxes.

Equation (9) shows that consumer prices depend on three components: *ad quantum* taxes, marginal costs gross of *ad valorem* taxation, and markups. The markup term depends on the demand system and therefore on substitution across legal brands and toward illicit alternatives. A lower effective illicit price makes legal residual demand more elastic and can constrain legal markups and pass-through. Stronger enforcement, higher search costs, or lower illicit availability can have the opposite effect.

This pricing block clarifies why excise evaluation cannot rely on statutory tax rates alone. A tax reform affects consumer prices through firms' pricing decisions; consumers respond to the resulting price vector; and fiscal revenue depends on the legal tax base that remains after those adjustments. The form of the tax matters because *ad valorem* and *ad quantum*

components enter the pricing problem differently and may affect relative prices across brands or categories.

3.3 Total consumption, legal sales, and the taxed base

The central distinction for excise evaluation is between total cigarette consumption and tax-paid legal sales. Let

$$Q_t^L = \sum_{n=1}^{N} q_{n,t},$$

denote legal tax-paid cigarette consumption, let $Q_t^I = q_t^I$ denote illicit cigarette consumption. Total cigarette consumption is therefore:

$$Q_t = Q_t^L + Q_t^I.$$

Public-health claims are naturally linked to Q_t, while fiscal outcomes depend on Q_t^L. These objects coincide only when illicit consumption is zero or when the legal share is constant and close to one. More generally, define the legal share of total cigarette consumption as

$$\theta_t = \frac{Q_t^L}{Q_t}, \quad 0 \leq \theta_t \leq 1.$$

Differentiating this expression we obtain:

$$dQ_t^L = \theta_t dQ_t + Q_t d\theta_t. \tag{10}$$

Equation (10) is central. A decline in legal sales can reflect a decline in total consumption, a decline in the legal share, or both. Thus, a fall in Q_t^L does not identify the sign of a change in Q_t. Total consumption may fall less than legal sales, may remain unchanged, or may even rise if substitution toward illicit products is strong enough. This is why legal sales cannot be interpreted mechanically as total cigarette consumption.

Fiscal revenue is collected only on legal tax-paid sales. From (8), the excise payment per unit of brand n is

$$T_{n,t} = \frac{\gamma_{n,t} p_{n,t} + \tau_{n,t}}{1 + \gamma_{n,t}}.$$

Total excise revenue is therefore:

$$R_t = \sum_{n=1}^{N} T_{n,t} q_{n,t},$$

which depends on the tax payment per legal unit, the size of legal sales, and the composition of legal sales across brands with different tax payments per unit. An increase in statutory taxation therefore does not guarantee an increase in revenue if the legal base contracts sufficiently or if sales move toward products with lower effective tax payments.

The joint claim requires two conditions. First, total cigarette consumption must fall. Second, fiscal revenue must rise. Neither condition follows mechanically from an increase in statutory taxation. The first depends on the response of total consumption, including substitution across legal and illicit products. The second depends on the response of the legal tax base and on the composition of legal sales across brands with different tax payments per unit. If the legal base is sufficiently elastic, revenue may fall even when the statutory tax burden rises. If substitution toward illicit cigarettes is sufficiently strong, legal sales may fall without a comparable reduction in total smoking.

This framework also clarifies the role of aggregation. An aggregate cigarette elasticity may combine several distinct margins: changes in total consumption, substitution across legal brands, movement across price segments, and substitution toward illicit products. For fiscal analysis, the relevant object is not simply the elasticity of aggregate cigarette purchases with respect to an average price. It is the response of tax-paid legal quantities and revenue to the effective price vector generated by the tax reform.

3.4 Welfare accounting

The "magic bullet" criterion is not itself a welfare criterion. A reform may reduce total smoking and increase revenue, but its welfare effect still depends on private surplus losses, externalities, internalities if admitted, enforcement costs, illicit-market responses, and the social value of public funds. Conversely, a reform may fail to raise revenue and still improve welfare if the externality or internality gains associated with lower total smoking

are sufficiently large. Welfare evaluation therefore requires separating the objects that are observed from those that must be assumed or inferred.

Let the change in private consumer surplus from cigarette consumption be decomposed as

$$\Delta CS = \Delta CS^L + \Delta CS^I,$$

Where ΔCS^L denotes the change in consumer surplus associated with legal cigarette consumption and ΔCS^I denotes the corresponding change associated with illicit or untaxed cigarette consumption. Let ΔR denote the change in excise revenue, η the social value of public funds, ν the per-unit uninternalized harm from cigarette consumption, ΔQ the change in total cigarette consumption, ΔQ^I the change in illicit consumption, κ the social cost per unit of illicit-market expansion, and Δe the change in enforcement or administrative costs. A first-pass welfare accounting expression is then:

$$\Delta W \approx \Delta CS^L + \Delta CS^I + \eta \Delta R - \nu \Delta Q - \kappa \Delta Q^I - \Delta e. \tag{11}$$

This expression makes the welfare logic explicit. The term $\Delta CS^L + \Delta CS^I$ captures the private surplus change from cigarette consumption. The term $\eta \Delta R$values the fiscal effect. The term $-\nu \Delta Q$ captures the externality or internality gain associated with any reduction in total smoking. Under a non-paternalistic benchmark, ν includes only harms not internalized by smokers, such as external costs. Under a behavioral benchmark, ν may also include internalities. The terms $\kappa \Delta Q^I$ and Δe capture costs associated with illicit-market expansion and enforcement.

Equation (11) clarifies why lower legal sales are not a welfare result. If legal sales fall because total consumption falls, the reform may generate externality or internality gains. If legal sales fall mainly because consumption shifts toward untaxed products, those gains may be small or absent, while revenue losses, illicit-market costs, and enforcement costs may become more important. Even a reduction in total consumption is not automatically welfare-improving: the associated externality or internality gains must be large enough to compensate for lost private surplus and other costs.

For $\Delta Q < 0$, equation (11) implies a break-even value of the uninternalized harm:

$$\nu^{*} = \frac{\Delta CS^{L} + \Delta CS^{I} + \eta \Delta R - \kappa \Delta Q^{I} - \Delta e}{\Delta Q}. \tag{12}$$

If $\nu > \nu^{*}$, the reform improves welfare under the assumptions used to construct the accounting exercise. If $\nu < \nu^{*}$, it does not. If $\Delta Q \geq 0$, the externality or internality channel cannot justify the reform through reduced total smoking; any welfare defense must come from other channels.

Equation (12) also clarifies the empirical limits of welfare evaluation. Some terms can be obtained or approximated from the available data, such as legal prices, legal quantities, legal tax payments, changes in real excise revenue, and the legal-market component of consumer surplus. Other terms can only be constructed under explicit assumptions, such as total consumption and illicit consumption, which depend on the unobserved legal share of the market. Finally, several welfare-relevant objects are not observed and cannot be estimated with the information available here, including ΔCS^{I}, ν, κ, Δe, and η. Any exercise that seeks to evaluate welfare must therefore state the assumptions imposed on these terms.

4 Bolivia: institutional setting, tax reforms, and data

Bolivia is a useful setting for evaluating tobacco excises because legal and untaxed cigarette markets coexist, the excise regime changed during the sample period, and the available data make it possible to observe a large part of the taxed legal market at monthly brand-level frequency. Evidence from Arteaga et al. (2024) and OECD (2024) indicates that illicit cigarettes are present and vary across cities, making the distinction between total consumption and tax-paid legal sales empirically relevant. The 2018 shift from *ad valorem* to *ad quantum* taxation also provides a concrete policy change through which effective tax burdens and relative prices may have changed across legal brands. This section describes the institutional and market setting, the evolution of the excise regime, and the data used in the empirical analysis.

4.1 Institutional setting and industry background

Bolivia's cigarette market operates in an institutional environment in which legal and illicit supply channels coexist. The country is also characterized by a high degree of informality. According to ILOSTAT, the informal employment rate in Bolivia was 83.9 percent in 2024. In such an environment, the reach of excise taxation is unlikely to depend only on statutory provisions. Enforcement, monitoring, distribution networks, and border control may shape the relative importance of legal and illicit channels and, therefore, the response of the taxed legal base.

Available evidence suggests that the cigarette market should not be treated as homogeneous. Arteaga et al. (2024), using smoker surveys and pack photography in La Paz/El Alto, Cochabamba, and Santa Cruz, report that 36.9 percent of cigarettes sold by unit or pack in their sample were illicit. They also report substantial heterogeneity across cities, with estimated illicit shares of 10.2 percent in La Paz/El Alto, 14.1 percent in Cochabamba, and 54.2 percent in Santa Cruz. According to the same study, illicit cigarettes in the sample were concentrated in a small number of Paraguayan brands, and Paraguay was the main origin of the illicit products identified.

These figures are not used as direct measurements of illicit quantities in the empirical analysis. They are nevertheless informative about the market environment. They indicate that legal cigarettes compete not only with other legal brands, but also with untaxed alternatives whose importance may vary across regions. This matters because the response of tax-paid legal sales to excise reform need not coincide with the response of total cigarette consumption.

The observed demand for cigarettes in Bolivia should also be interpreted in light of the broader regulatory environment. Arteaga et al. (2024) identify three main pillars of the current framework: the 1986 tax reform, the 2005 ratification of the Framework Convention on Tobacco Control, and the 2020 law on prevention and control of tobacco use. The legal framework includes restrictions on smoking in enclosed public and workplace spaces, warning-label requirements, and prohibitions on some forms of tobacco advertising

and promotion. These measures may affect cigarette demand independently of taxation and are therefore relevant when interpreting the evolution of legal sales.

4.2 Evolution of excise policy

The taxation of cigarettes in Bolivia originates in the 1986 tax reform, under which cigarettes became subject to the specific consumption tax, known as ICE (*Impuesto a los Consumos Específicos*), initially applied on an *ad valorem* basis over factory price (Arteaga et al., 2024). Two later changes are especially relevant for this paper. First, in 2010, the *ad valorem* rate was increased for cigarettes made with blond tobacco. Second, a reform implemented in 2018 replaced the *ad valorem* cigarette excise with an *ad quantum* regime (Arteaga et al., 2024; OECD, 2024).

The 2018 reform is central because it changed the form of taxation. Under the previous regime, tax liability depended on declared value or factory price. Under the new regime, tax liability depends on physical quantities, with nominal amounts indexed through the Housing Development Unit (*Unidad de Fomento a la Vivienda*, UFV) and adjusted annually.

The reform should not be interpreted as the introduction of a single uniform per-unit tax across all cigarettes. The *ad quantum* amounts differ between dark and blond cigarettes. Therefore, the effective tax share must be interpreted within product categories as well as across them. Within a given category-specific unit tax, lower-price brands mechanically face a higher effective tax share than higher-price brands because the tax is independent of price. This distinction is important for the empirical analysis because the paper studies how the tax reform interacted with relative prices, brand composition, legal sales, and the taxed base.

Arteaga et al. (2024) and Morales and Santander (2025) also indicate that, despite these reforms, Bolivia's overall statutory tax burden remained below benchmarks commonly used in international comparisons. For the purposes of this paper, however, the relevant issue is not only whether Bolivia's average tax burden was low or high relative to other

countries. The relevant issue is how the reform changed the structure of taxation in a differentiated market where consumers may substitute across legal brands and toward untaxed alternatives.

4.3 Data

The empirical analysis uses a monthly panel of legal cigarette sales and prices covering the period from April 2015 to December 2025. The core dataset was provided by Compañía Industrial de Tabacos S.A. (CITSA) and contains observations for the cigarette brands sold by the company in each of Bolivia's nine departments. The available variables include quantity sold, measured in sticks; gross sales revenue; net revenue, defined as gross revenue net of ICE payments; and ICE payments. Gross prices are constructed as gross sales revenue divided by quantity sold. Net-of-ICE prices are constructed as net revenue divided by quantity sold.

CITSA represents a large and relatively stable share of the legal cigarette market in Bolivia. According to the information available for this study, its average market share over the period considered was approximately 77 percent of legal cigarette sales. This makes the dataset unusually informative about the evolution of the formal taxed market. At the same time, the data do not cover the entire legal market and do not observe illicit sales or total cigarette consumption. The empirical object directly observed in the analysis is therefore CITSA's legal segment, which is a sizable part of the formal market but not the universe of cigarette consumption in Bolivia. This limitation is central to the paper's interpretation: observed legal sales, even from a large legal supplier, cannot be treated as total cigarette consumption.

The paper uses anonymized brand labels. These labels preserve the distinction between dark and blond cigarettes, because that distinction is relevant for the tax regime and for the empirical analysis. The three dark-tobacco brands included in the econometric analysis are denoted DK1, DK2, and DK3. The four blond-tobacco brands included in the econometric analysis are denoted BL1, BL2, BL3, and BL4. These labels are generic and do not reproduce commercial brand names.

Three additional blond-tobacco brands are excluded from the econometric exercises because their time coverage is not comparable to that of the persistent brands. One brand was sold only until 2019, another is observed from 2017 to 2024, and a third began to be sold only in 2024. These brands are relevant for describing the broader product set, but their incomplete and uneven coverage makes them less suitable for brand-level demand estimation over the full sample. The empirical results should therefore be interpreted as estimates for the persistent observed legal brands, not for every brand sold during the period.

The CITSA panel is complemented with information from the National Institute of Statistics (INE, Instituto Nacional de Estadística) on the monthly price of the "cigarette" category used in the construction of the consumer price index in each department. Because the precise composition of this category cannot be matched to specific brands, this series is not interpreted as the price of any individual competing brand. It is used instead as a proxy for broader cigarette-price conditions in each department and month. This distinction is important because aggregate cigarette-price measures may combine true price variation with changes in brand composition.

To capture competitive pressure from external and potentially illicit sources, the dataset also includes monthly observations on the price in Paraguay of a representative cigarette product. Paraguayan prices are converted into bolivianos using monthly exchange-rate information. This variable is used as a proxy for changes in the relative attractiveness of an important external source of cigarette supply and, more generally, for competitive pressure from Paraguayan-origin products.

The analysis also includes department-level monthly prices for goods that may plausibly act as complements or substitutes for cigarette consumption: coffee, beer, coca leaf used for chewing, and chewing gum. These series are obtained from INE. Their inclusion is motivated by the demand framework in Section 3, in which cigarette consumption may depend not only on own prices and cigarette substitutes, but also on the prices of related goods entering the consumer's choice problem.

Income and cost conditions are captured with additional INE series. The monthly average national wage rate is used as a proxy for income. On the cost side, the analysis includes the average wage in the tobacco industry as a cost shifter relevant for legal cigarette supply. A second cost shifter is the monthly price of electricity by department. These variables are used to capture changes in production or distribution costs that may affect legal cigarette prices independently of contemporaneous demand shocks.

Brand-level prices, the representative cigarette-category price, the Paraguayan cigarette price, and the prices of related goods are expressed in real terms using the department-level consumer price index reported by INE. Wage variables are also deflated to obtain real measures of income and costs.

The monthly panel is complemented with annual firm-level information from CITSA on formal-sector outcomes. These include production, employment, labor costs, and purchases of dark tobacco from local farmers. These annual series are not used to estimate demand. They are used to evaluate whether the deterioration of the taxed legal segment was accompanied by changes upstream and downstream in the formal production chain.

4.4 Empirical counterparts and limitations

The data provide direct information on legal brand-level quantities, gross prices, net-of-ICE prices, and tax payments. They also provide proxies for broader cigarette-price conditions, external competitive pressure, related goods, income, and selected cost shifters. This structure allows the empirical analysis to connect the data to the framework in Section 3 along four margins.

First, brand-level prices and quantities are used to estimate legal demand separately by brand. This preserves heterogeneity in price responses and avoids imposing a common elasticity across products that differ in price segment, tobacco type, and exposure to the tax regime. The relevant empirical object is legal tax-paid demand in the observed segment, not total cigarette consumption.

Second, gross prices, net-of-ICE prices, and tax payments allow the analysis to study effective tax burdens and price adjustment across brands. This is the empirical counterpart of the pricing block in Section 3.2. The data make it possible to observe how the tax component, the final consumer price, and the net-of-tax price evolved, but they do not by themselves identify a full structural pass-through model.

Third, ICE payments and legal quantities allow the paper to study the evolution of the taxed legal base and real excise revenue. This is the fiscal object relevant for the "magic bullet" claim. Unlike total consumption, this margin is directly observed in the legal segment covered by the data.

Fourth, the annual production-chain indicators provide evidence on whether the contraction of the legal taxed segment was accompanied by broader changes in the formal sector. These indicators are not used to estimate demand. They are used to assess whether the evolution of legal sales and revenue coincided with changes upstream and downstream in the formal production chain.

The main limitation is that the data do not directly observe total cigarette consumption, illicit consumption, or the full legal market outside CITSA. The paper therefore does not estimate the total-consumption response to the tax reform, nor does it estimate the size of the illicit market. Its empirical claims are more limited but still policy-relevant: legal sales cannot be treated as total smoking, aggregate legal demand can conceal important substitution margins, and the fiscal performance of an excise reform must be evaluated using the taxed legal base and observed revenue, not statutory tax rates alone.

5 Empirical analysis

This section applies the framework of Section 3 to the Bolivian legal cigarette market. The analysis focuses on the legal segment observed in the CITSA data. This segment represents a large share of legal cigarette sales in Bolivia, but it is not the universe of cigarette consumption. The results should therefore be interpreted as evidence on a sizable part of

the taxed legal market, not as direct evidence on total smoking or the full size of the illicit market.

The analysis proceeds in six steps. Section 5.1 documents the aggregate evolution of legal sales, real revenue, real ICE payments, and average real prices in the observed legal segment. Section 5.2 estimates legal cigarette demand, distinguishing a static conditional benchmark, a lagged-demand specification, and a forward-looking rational-addiction exercise. Section 5.3 examines effective tax burdens and price adjustment, using the evolution of gross prices, net-of-tax prices, and tax payments across brands. Section 5.4 studies the legal production chain, assessing whether the contraction of the taxed legal segment was accompanied by changes in production, employment, labor costs, and purchases from domestic suppliers. Section 5.5 returns to the distinction between total consumption and the taxed legal base. Section 5.6 discusses welfare implications using the accounting framework developed above.

5.1 Aggregate trends in the observed legal segment

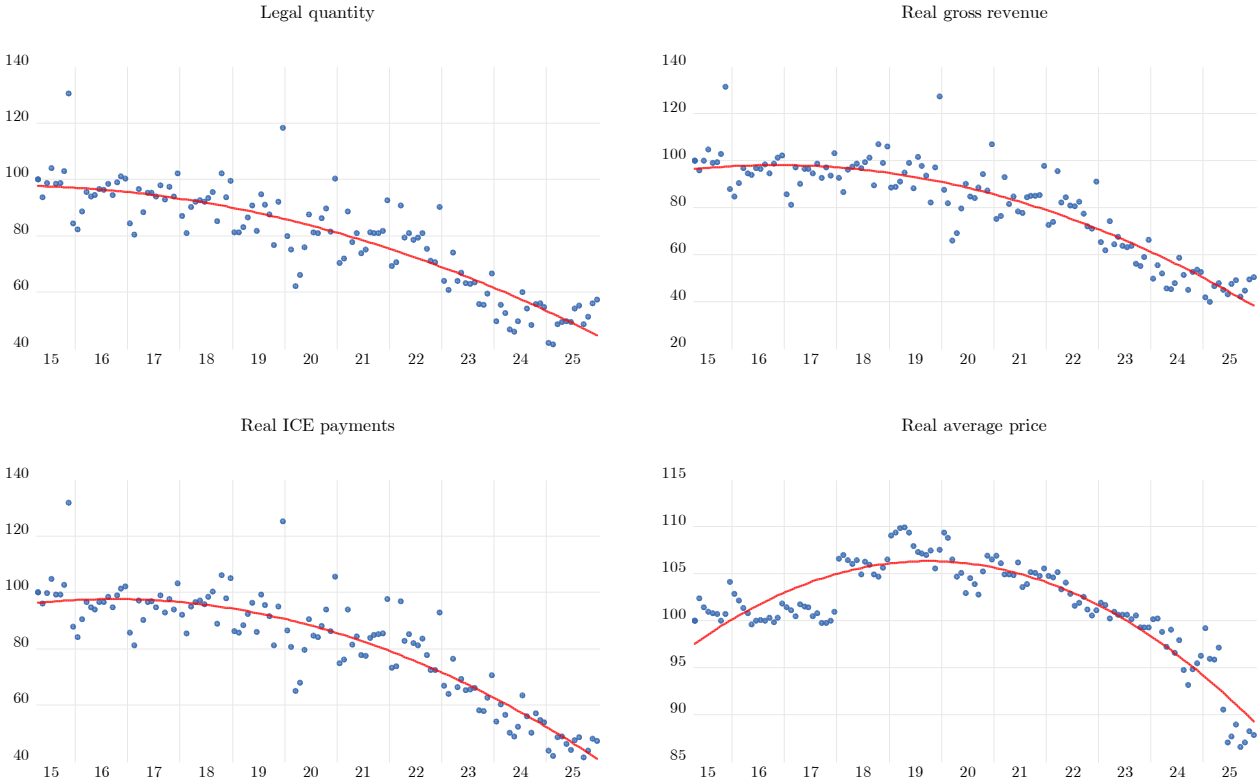
Before turning to the econometric analysis, it is useful to summarize the aggregate evolution of the observed legal segment. Figure 1 reports indexed series for total legal quantity sold, real gross revenue, real ICE payments, and the real average final price, with April 2015 normalized to 100.

The real average final price is constructed as total real gross revenue divided by total quantity sold. Equivalently, it is a sales-weighted average of brand-level prices, with weights given by each brand's share in total legal sales. This distinction matters. The aggregate price is not a fixed-price index for a homogeneous product. Like the unit values commonly used in household expenditure surveys, it combines changes in within-brand prices with changes in the composition of purchases across brands. It is therefore useful as a descriptive summary of the legal market, but it should not be interpreted mechanically as an exogenous market price.

Three patterns stand out. First, legal sales declined substantially over the sample period. By the end of the sample, the quantity index is close to one half of its initial level. Second, real gross revenue and real ICE payments declined even more sharply. Third, the real average final price increased after the replacement of the *ad valorem* ICE by an *ad quantum* regime in 2018, but that increase was not sustained. By the end of the sample, the real average price was below its initial level.

Figure 1

Indexed legal quantity, gross revenue, ICE payments, and average final price



Notes: All panels report indexed series normalized to 100 in April 2015, the initial observation in the sample. Quantity corresponds to total legal sticks sold in the observed legal segment. Real gross revenue, real ICE payments, and the real average final price are deflated using the national consumer price index (CPI). The real average final price is constructed as real gross revenue divided by quantity sold and is therefore a sales-weighted average of brand-level prices. Solid fitted lines are obtained from quadratic regressions on a time trend and are included to summarize low-frequency movements.

These patterns are informative for the "magic bullet" claim. In the observed legal segment, legal sales fell, but real excise revenue also fell. Thus, the strong fiscal version of the claim

did not materialize in the legal market observed in the data. The decline in legal sales cannot be interpreted as an equivalent decline in total cigarette consumption, because total consumption is not observed and because part of the adjustment may have occurred through substitution across legal brands or toward untaxed alternatives. Still, the fiscal object is directly observed: the legal tax base contracted enough for real ICE revenue to fall.

Figure 1 also provides a first indication of why aggregate evidence must be interpreted cautiously. Real gross revenue and real ICE payments fell by more than legal quantity because the contraction of the legal market was accompanied by a weak real-price path. But the aggregate real price itself is affected by changes in brand composition. Thus, the figure is best read as a descriptive summary of the observed legal segment, not as evidence from a single representative cigarette market. This motivates the brand-level analysis below, which separates own-price responses from substitution across products and from the compositional movements embedded in aggregate unit values.

5.2 Legal cigarette demand

The demand framework in Section 3 motivates three empirical specifications. The first is a static conditional demand equation, which serves as the main benchmark. The second is a lagged-demand specification that allows legal purchases to display persistence and permits short-run and long-run elasticities to differ. The third is a forward-looking rational-addiction exercise, which follows the Euler-equation logic of Section 3.1 by including future consumption and future prices. The static and lagged-demand specifications provide the main quantitative evidence. The forward-looking model is theoretically relevant, but empirically demanding; with the available instruments, it is not identified with enough precision to provide reliable policy elasticities.

All demand equations are estimated separately by brand. This is important because the central empirical question is whether aggregation hides economically relevant substitution margins. A pooled specification imposing common price elasticities across brands would obscure precisely the heterogeneity that matters for evaluating the legal tax base.

Aggregate estimates are reported only as a comparison and should be interpreted cautiously because the aggregate price is a sales-weighted average of brand-level prices. As discussed above, this object shares the same limitation as unit values in household surveys: it combines within-brand price movements with changes in brand composition.

Throughout this subsection, the dependent variable is legal tax-paid sales in the observed CITSA segment. The estimates therefore describe legal demand in the observed market, not total cigarette consumption. This distinction is essential for interpreting the results: brand-level elasticities are informative about the vulnerability of the taxed legal base, but they do not by themselves identify the response of total smoking.

5.2.1 Static conditional demand benchmark

The first specification is a static conditional demand equation estimated separately by brand. The model abstracts from transitional dynamics and can therefore be interpreted as a conditional long-run benchmark: it relates legal sales to prices and controls after short-run adjustment margins are not modeled explicitly. This interpretation is useful because the policy question concerns the sustained response of the legal tax base, not only the immediate response to a monthly price movement. At the same time, the specification should not be read as a full structural dynamic model. It estimates conditional long-run associations between legal sales, prices, and controls, while the dynamic specifications below examine whether allowing for habit persistence or forward-looking behavior materially changes the conclusions.

For each brand n, the baseline static specification is therefore written as:

$$\begin{aligned}\ln q_{n,d,t} = \alpha + \beta_1 \ln p_{n,d,t} + \beta_2 \ln P_{d,t} + \beta_3 \ln p_t^I + \phi_n{}' w_{d,t} + \beta_4 \ln y_t \\ +\gamma_d + \lambda_{d,year} + \delta_{month,t} + u_{n,d,t},\end{aligned} \tag{13}$$

where $q_{n,d,t}$ denotes legal sales of brand n in department d and period t, $p_{n,d,t}$ is the own real price of the brand, $P_{d,t}$ is the real INE cigarette-category price, p_t^I is the real local-currency price of the representative Paraguayan cigarette, $w_{d,t}$ is the vector of log real prices of the other goods considered (coffee, beer, coca leaf, and chewing gum), y_t is the

(log) real wage measure, γ_d denotes department fixed effects, $\lambda_{d,year}$ department-specific year effects, and $\delta_{month,t}$ seasonal dummies.

The interpretation of $P_{d,t}$ requires some care. It is not the price of a specific competing brand. It is a broader cigarette-category price constructed by the statistical agency and is used here as a proxy for general cigarette-market price conditions in each department and month. The Paraguayan price proxy captures changes in the relative attractiveness of an important external source of cigarette supply and, more generally, competitive pressure from Paraguayan-origin products. The prices of coffee, beer, coca leaf, and chewing gum allow cigarette demand to depend on other goods that may operate as complements or substitutes.

Inference is based on wild cluster bootstrap procedures at the department level. This choice is dictated by the small number of clusters: the panel contains only nine departments, so conventional cluster-robust asymptotic approximations may provide unreliable inference. Wild cluster bootstrap methods are commonly recommended in this setting because they improve finite-sample inference with clustered errors relative to standard cluster-robust tests (Cameron et al., 2008). The procedure should nevertheless be interpreted with the usual caution, especially when the effective number of clusters is small, leverage is high, or cluster sizes are unbalanced (MacKinnon et al., 2023). The tables therefore report the preferred bootstrap-based standard errors.

Appendix A addresses the main endogeneity concern: the possible endogeneity of the brand's own price. Prices and quantities are equilibrium outcomes, so demand shocks may induce price adjustments. Appendix A therefore reports instrumental-variable estimates that treat $p_{n,d,t}$ as endogenous, using supply-side cost shifters and lagged market variables as excluded instruments. The Durbin–Wu–Hausman tests do not provide systematic evidence against treating own prices as exogenous in the conditional demand equations, and the Hansen overidentifying-restriction tests generally fail to reject the validity of the overidentifying restrictions. The main limitation of the IV exercise is weak identification: the Lewis and Mertens (2025) weak-instrument test fails to reject weak identification in all specifications except DK3. As a result, the IV estimates do not provide a stable or

better-identified alternative to the static conditional benchmark. In Hendry's (1995) terminology, the preferred interpretation is one of weak exogeneity for the parameters of the conditional legal-demand equation, not structural exogeneity of prices.

A secondary concern is the broader cigarette-category price proxy, $P_{d,t}$. Specifications treating both $p_{n,d,t}$ and $P_{d,t}$ as endogenous were also examined, but they were more unstable and did not materially change the conclusions. They are therefore not reported.

Appendix B provides an additional robustness check by including a COVID-period control, defined as a dummy equal to one from April to December 2020. The own-price estimates are very close to the baseline results, and the aggregate elasticity remains small relative to the brand-level elasticities. Thus, the main demand results are not driven by the inclusion of the main pandemic-disruption period.

Table 1

OLS estimates of legal cigarette demand by brand and in the aggregate

Variable	DK1	DK2	DK3	BL1	BL2	BL3	BL4	Total
$p_{n,d,t}$	-1.97 (0.78)	-5.57 (1.22)	-2.13 (0.72)	-0.68 (0.49)	-1.15 (0.72)	-2.21 (0.41)	-1.93 (0.95)	-0.22 (0.27)
$P_{d,t}$	0.66 (0.39)	0.66 (0.54)	0.42 (0.26)	0.27 (0.34)	0.31 (0.07)	-0.50 (0.31)	0.53 (0.31)	0.26 (0.12)
p_t^I	0.44 (0.12)	1.39 (0.44)	0.43 (0.22)	0.48 (0.15)	0.30 (0.20)	1.74 (0.51)	-0.63 (0.37)	0.15 (0.14)
$p_{d,t}^{coffee}$	0.29 (0.17)	-0.48 (0.31)	0.04 (0.08)	-0.05 (0.11)	-0.05 (0.11)	-0.41 (0.18)	-0.26 (0.24)	0.01 (0.06)
$p_{d,t}^{beer}$	0.64 (0.24)	0.20 (0.43)	0.64 (0.23)	-0.60 (0.62)	-0.18 (0.17)	-1.67 (0.83)	0.51 (0.97)	0.00 (0.16)
$p_{d,t}^{coca}$	-0.27 (0.09)	-0.45 (0.16)	-0.21 (0.05)	-0.03 (0.04)	-0.01 (0.06)	0.05 (0.24)	-0.28 (0.10)	-0.05 (0.05)
$p_{d,t}^{gum}$	-0.18 (0.19)	0.03 (0.16)	0.02 (0.10)	-0.06 (0.06)	-0.02 (0.06)	-0.44 (0.39)	0.46 (0.22)	-0.06 (0.04)
y_t	1.34 (0.48)	2.68 (2.49)	-0.42 (0.60)	0.63 (0.45)	0.57 (0.66)	2.51 (1.10)	-1.44 (0.79)	-0.39 (0.31)
Observations	1161	1107	1161	1161	1161	1156	1161	1161
Adjusted R^2	0.98	0.96	0.98	0.99	0.98	0.95	0.96	0.99
SER	0.28	0.47	0.18	0.17	0.15	0.36	0.35	0.12
DW	2.33	1.93	2.09	2.10	1.89	1.92	2.15	1.93

Notes: OLS estimates of the log-linear static demand equations. Dependent variable: log legal quantity sold. Continuous regressors are in logs and, where applicable, real terms. All specifications include monthly seasonal dummies, department fixed effects, and department-specific year effects. Wild cluster bootstrap standard errors, clustered by department, in parentheses. The Total column refers to aggregate legal sales in the observed segment; its price variable is the sales-weighted average final price. SER: standard error of the regression; DW: Durbin–Watson statistic.

Table 1 reports the static estimates. The own-price coefficient is negative in all brand-level equations and in the aggregate. The central result is the contrast between brand-level and aggregate elasticities. At the brand level, legal demand is generally much more price-responsive than at the aggregate level. In the aggregate equation, the own-price elasticity is small. This difference is not a minor empirical detail. It indicates that the aggregate equation is not simply a noisier version of the brand-level equations; it answers a different question.

These brand-level magnitudes should be read accordingly. Most exceed unity in absolute value and DK2 reaches about -5.6—large relative to the aggregate cigarette elasticities common in the literature, but not the same object: a brand-level own-price elasticity absorbs substitution toward other legal brands and toward untaxed alternatives. In a differentiated market a brand can shed much of its volume to close substitutes after a relative-price change while total smoking moves little, so large brand-level elasticities are the expected signature of substitution rather than implausible demand responses, and they speak to the responsiveness of the taxed legal base.

The aggregate specification is included only as a comparison. Its price variable is the sales-weighted average price discussed in Section 5.1, which combines within-brand price changes with changes in brand composition. It shares the same conceptual limitation as unit values constructed from household expenditure surveys. Appendix C reinforces this interpretation. The variance decomposition shows that, at the departmental level, changes in aggregate legal unit values are driven more by the brand-composition component than by the within-brand price component, often by a wide margin. National aggregation compresses this distinction. Once prices are collapsed into an aggregate unit value, the data lose much of their ability to reveal how important brand reallocation actually is.

This point has broader implications for elasticity estimates based on household expenditure surveys. In that literature, the relevant "price" is typically a unit value, constructed as expenditure divided by quantity, without systematically distinguishing brands and often without properly observing illicit purchases. In a market with substantial

brand substitution and illicit trade, variation in such unit values is unlikely to reflect only common price variation. It may instead combine changes in legal brand composition, access to illicit cigarettes, household quality choice, and genuine market-wide price movements. For cigarettes, this is particularly problematic because what appears as "quality variation" is often not a gradual margin, but discrete movement across brands, including movement toward cheaper or illicit products. If so, estimates based on aggregate unit values may be poor guides to the policy-relevant elasticity for excise reform.

The central implication is therefore not simply that the aggregate elasticity is small, but that the aggregate object being estimated may be the wrong one for policy. For excise evaluation, the relevant object is not the elasticity of aggregate legal sales with respect to an average price, but the responsiveness of the taxed base once substitution across brands, across segments, and toward untaxed products is allowed. In that sense, aggregation may understate the contraction of the legal tax base and bias the analysis toward an overly optimistic view of excise reforms.

The prices of other candidate complements or substitutes play a less systematic role. Coffee, beer, and chewing gum do not display stable patterns across brands. Coca leaf tends to enter with a more consistently negative sign in some specifications, suggesting complementarity for some segments, but this is not the main driver of the results. Income effects are also limited and heterogeneous across brands.

5.2.2 Lagged-demand dynamics

The static benchmark abstracts from explicit dynamic adjustment. To evaluate whether this omission drives the main results, the next specification adds lagged legal purchases to equation (13). This provides a reduced-form way of capturing persistence in legal demand and allows short-run and long-run price elasticities to differ. The lagged term should not be interpreted mechanically as evidence of addiction. It may reflect habit persistence, but it may also capture gradual adjustment, inventory behavior, persistent unobserved heterogeneity, or slow substitution across products and channels.

For each brand n, the specification is

$$\begin{aligned}\ln q_{n,d,t} = \alpha + \beta_1 \ln p_{n,d,t} + \beta_2 \ln P_{d,t} + \beta_3 \ln p_t^I + \phi_n{}' w_{d,t} + \beta_4 \ln y_t \\ +\beta_5 \ln q_{n,d,t-1} + \gamma_d + \lambda_{d,year} + \delta_{month,t} + u_{n,d,t}.\end{aligned} \tag{14}$$

The short-run own price elasticity is β_1. When $\left|\beta_5\right| < 1$, the implied long-run own-price elasticity is

$$\varepsilon_n = \frac{\beta_1}{1-\beta_5}. \tag{15}$$

This long-run elasticity measures the cumulative response of legal tax-paid sales within the reduced-form lagged-demand model. It is not a structural estimate of the long-run response of total cigarette consumption. If $\beta_5 > 0$, the long-run elasticity is larger in absolute value than the short-run elasticity. If β_5 is close to zero, the short-run and long-run elasticities are similar. If $\beta_5 < 0$, the dynamic response partly offsets the initial effect, which may reflect inventory behavior, short-run displacement across months, promotional timing, measurement noise, or other forms of adjustment that are not literal habit formation.

Inference for equation (14) requires a bootstrap design that respects the dynamic structure of the model. A static residual perturbation of the type used for equation (13) would generate pseudo-samples in which the lagged dependent variable is no longer internally consistent with the simulated dependent variable. The lagged-demand specification is therefore estimated using a recursive-design wild cluster bootstrap. After estimating equation (14) by OLS, the coefficient vector, including the estimated persistence parameter $\hat{\beta}_5$, and the empirical residuals are retained. In each bootstrap replication, a single two-point wild-bootstrap weight of the type proposed by Mammen (1993) is drawn for each of the nine departments and applied to all residuals in that department. This preserves arbitrary intra-departmental correlation while allowing residual perturbations to vary across clusters, as in the wild cluster bootstrap logic of Cameron et al. (2008).

The pseudo-sample is then reconstructed recursively. For each department-brand series, the first observation is initialized using the observed lagged value. Subsequent simulated log quantities are generated using the estimated coefficients and the bootstrap residuals,

replacing the lagged dependent variable at each step with the simulated value from the previous period. Equation (14) is then re-estimated on the simulated sample. Standard errors are computed as the cross-replication standard deviations of the bootstrap estimates. For the long-run elasticity, the nonlinear transformation in (15) is recomputed in each bootstrap replication. This procedure combines the small-cluster motivation of Cameron et al. (2008) with the recursive-design wild bootstrap for autoregressive models with conditional heteroskedasticity studied by Gonçalves and Kilian (2004).

Table 2

Lagged-demand estimates of legal cigarette demand

Variable	DK1	DK2	DK3	BL1	BL2	BL3	BL4	Total
$p_{n,d,t}$	-2.11 (0.46)	-2.51 (2.46)	-2.07 (0.50)	-0.67 (0.24)	-1.14 (0.42)	-2.27 (0.66)	-1.96 (0.36)	-0.25 (0.19)
$P_{d,t}$	0.74 (0.42)	-0.58 (0.60)	0.43 (0.24)	0.25 (0.18)	0.32 (0.13)	-0.42 (0.37)	0.45 (0.38)	0.26 (0.13)
p_t^f	0.39 (0.26)	1.25 (0.62)	0.42 (0.19)	0.46 (0.20)	0.28 (0.15)	1.37 (0.57)	-0.86 (0.37)	0.12 (0.12)
$p_{d,t}^{coffee}$	0.40 (0.23)	-0.03 (0.28)	0.05 (0.08)	-0.05 (0.11)	-0.07 (0.10)	-0.36 (0.16)	-0.26 (0.21)	-0.01 (0.06)
$p_{d,t}^{beer}$	0.71 (0.37)	0.27 (0.44)	0.65 (0.30)	-0.61 (0.57)	-0.18 (0.17)	-1.61 (0.88)	0.53 (0.67)	-0.00 (0.17)
$p_{d,t}^{coca}$	-0.34 (0.15)	-0.33 (0.23)	-0.23 (0.09)	-0.02 (0.04)	-0.01 (0.07)	0.10 (0.27)	-0.28 (0.10)	-0.04 (0.06)
$p_{d,t}^{gum}$	-0.27 (0.23)	0.62 (0.54)	0.01 (0.10)	-0.06 (0.08)	-0.03 (0.06)	-0.31 (0.47)	0.50 (0.22)	-0.06 (0.05)
y_t	1.58 (0.76)	3.24 (2.36)	-0.57 (0.51)	0.74 (0.48)	0.53 (0.55)	3.03 (1.38)	-1.20 (0.44)	-0.34 (0.33)
$q_{n,d,t-1}$	-0.19 (0.02)	0.38 (0.08)	-0.07 (0.04)	-0.02 (0.02)	0.08 (0.02)	0.05 (0.03)	-0.03 (0.05)	0.07 (0.02)
Observations	1152	1077	1152	1152	1152	1146	1152	1152
Adjusted R^2	0.98	0.95	0.98	0.99	0.98	0.95	0.96	0.99
SER	0.28	0.51	0.18	0.17	0.15	0.36	0.35	0.12
DW	2.00	2.16	1.97	2.06	2.03	2.01	2.06	2.06

Notes: OLS estimates of the log-linear habit-persistence demand equations, which add the first lag of legal quantity sold. Dependent variable: log legal quantity sold. Continuous regressors are in logs and, where applicable, real terms. All specifications include monthly seasonal dummies, department fixed effects, and department-specific year effects. Standard errors in parentheses use the recursive-design wild cluster bootstrap, clustered by department. The Total column refers to aggregate legal sales in the observed segment; its price variable is the sales-weighted average final price. SER: standard error of the regression; DW: Durbin–Watson statistic.

Table 2 reports the lagged-demand estimates. The main conclusion from the static benchmark is preserved. Brand-level legal demand remains much more price-responsive than aggregate legal demand. Lagged legal purchases are statistically relevant in some cases, but the estimated persistence is not sufficiently large or systematic to overturn the

interpretation of the static estimates. Consequently, the implied long-run elasticities differ from the short-run elasticities, but the difference is generally not large enough to change the central conclusion.

The results are informative in two ways. First, the contrast between brand-level and aggregate elasticities is not an artifact of omitting simple persistence in legal purchases. Once lagged sales are included, the aggregate elasticity remains small, while several brand-level elasticities remain elastic. Second, the legal market appears to adjust importantly through relative-price reallocation rather than only through gradual dynamic adjustment. This reinforces the interpretation that the legal tax base is vulnerable because consumers can substitute across legal brands and possibly across channels, not merely because aggregate cigarette demand adjusts slowly over time.

The lagged-demand specification therefore strengthens the main empirical message without becoming the preferred structural model. It confirms that the static conditional benchmark is not misleading because of a simple omitted lagged-consumption term. At the same time, it remains a reduced-form dynamic specification. The next subsection considers the more demanding forward-looking model implied by the rational-addiction framework.

5.2.3 Forward-looking rational-addiction specification

The final demand specification follows the forward-looking logic of the rational-addiction framework in Section 3.1. Unlike the static and lagged-demand specifications, this model allows current legal purchases to depend not only on current prices and past consumption, but also on expected future consumption and expected future prices. This makes the specification closer to the Euler-equation interpretation of dynamic cigarette demand, but also considerably more demanding empirically.

For each brand n, the specification is

$$\begin{aligned}\ln q_{n,d,t} = \alpha + \beta_1 \ln p_{n,d,t} + \beta_2 \ln P_{d,t} + \beta_3 \ln p_t^I + \phi_n{}' w_{d,t} + \beta_4 \ln y_t \\ +\beta_5 \ln q_{n,d,t+1} + \beta_6 \ln p_{n,d,t+1} + \gamma_d + \lambda_{d,year} + \delta_{month,t} + u_{n,d,t}.\end{aligned} \tag{16}$$

The variables are defined as in equation (13). The coefficient β_5 captures the forward-looking consumption term, while β_1 and β_6 measure the effects of current and future own prices. The variables $\ln q_{n,d,t+1}$, $\ln p_{n,d,t}$, and $\ln p_{n,d,t+1}$ are treated as endogenous. The broader cigarette-category price $P_{d,t}$ is treated as weakly exogenous in the preferred implementation, consistent with the diagnostic evidence discussed in Appendix A.

The model is estimated by GMM using instruments available at time t. The instrument set includes supply-side cost shifters and lagged market variables. The objective is to exploit variation that predicts current and future legal prices, as well as future legal purchases, without being driven by contemporaneous structural demand shocks. This is a demanding requirement, so the rational-addiction specification also provides a useful diagnostic of how far the available data can support the richer dynamic model implied by the theory.

Inference for equation (16) uses a wild cluster bootstrap of the GMM estimator: unrestricted residuals are taken from the GMM fit, perturbed by department-level two-point wild weights (Mammen, 1993) with regressors and instruments held fixed, and the model is re-estimated by GMM in each replication, with standard errors equal to the cross-replication standard deviations.[2]

For a permanent price change in a steady state in which $q_{n,d,t+1} = q_{n,d,t}$ and $p_{n,d,t+1} = p_{n,d,t}$, the implied long-run own-price elasticity is

$$\varepsilon_n = \frac{\beta_1 + \beta_6}{1 - \beta_5}. \tag{17}$$

This expression differs from the habit-persistence elasticity in equation (15). In the habit-persistence model, long-run adjustment works through the lagged dependent variable. In

[2] The GMM rather than OLS residuals are used because OLS residuals would reintroduce the endogeneity and recenter the bootstrap on the OLS estimate (Davidson and MacKinnon, 2010; Cameron and Miller, 2015). No recursive reconstruction is needed here, unlike Section 5.2.2, since identification works through the moment conditions and the fixed instruments.

the rational-addiction specification, the permanent price effect combines current and future price coefficients and scales them by the forward-looking consumption coefficient. Reliable estimation of equation (17) therefore requires precise estimates of all three parameters.

Table 3

GMM estimates of the rational-addiction demand model

Variable	DK1	DK2	DK3	BL1	BL2	BL3	BL4	Total
$q_{n,d,t+1}$	2.60 (3.23)	1.11 (0.12)	0.59 (1.10)	1.22 (4.40)	-0.76 (0.66)	1.71 (1.17)	-2.15 (7.80)	0.02 (1.31)
$p_{n,d,t}$	-16.72 (28.5)	-7.39 (12.7)	-8.26 (7.54)	14.28 (61.3)	1.73 (2.52)	-2.76 (9.87)	-13.25 (31.5)	2.53 (2.42)
$p_{n,d,t+1}$	47.96 (71.8)	7.44 (14.9)	16.72 (12.7)	-13.57 (45.7)	-6.85 (7.56)	4.99 (12.1)	9.37 (24.1)	-3.66 (4.25)
P_t	-3.70 (5.70)	0.06 (0.22)	-0.21 (0.26)	0.01 (2.03)	0.57 (0.25)	0.82 (1.00)	2.72 (8.93)	0.33 (0.43)
p_t^I	-2.57 (4.67)	-0.69 (1.78)	-2.09 (1.81)	0.66 (0.52)	0.69 (0.73)	-0.26 (1.04)	-3.12 (8.47)	0.21 (0.75)
$p_{d,t}^{coffee}$	4.59 (6.67)	-0.07 (0.10)	1.04 (0.66)	0.17 (1.39)	-0.46 (0.50)	1.29 (1.32)	-0.45 (1.13)	-0.12 (0.40)
$p_{d,t}^{beer}$	-5.07 (8.31)	0.01 (0.16)	-0.98 (0.65)	-0.24 (1.00)	-0.14 (0.48)	0.03 (1.75)	2.44 (7.95)	-0.10 (0.39)
$p_{d,t}^{coca}$	-0.54 (0.39)	-0.02 (0.14)	-0.39 (0.14)	-0.05 (0.29)	0.12 (0.24)	-0.50 (0.53)	-0.95 (2.12)	-0.06 (0.16)
$p_{d,t}^{gum}$	-6.46 (10.2)	0.64 (1.17)	-1.20 (0.81)	-0.62 (3.53)	0.03 (0.20)	-0.43 (0.70)	3.37 (8.97)	-0.00 (0.29)
y_t	-3.08 (8.65)	1.13 (1.43)	-1.16 (2.55)	-2.66 (16.2)	0.36 (1.13)	0.41 (2.10)	-4.68 (14.7)	-0.93 (0.82)
Observations	1143	1055	1143	1143	1143	1137	1143	1143
Adjusted R^2	0.54	0.92	0.94	0.94	0.96	0.83	0.62	0.98
SER	1.28	0.64	0.35	0.53	0.23	0.68	1.02	0.16
DW	2.32	2.89	2.22	2.89	1.42	2.80	1.75	2.35
DWH	0.01	0.00	0.00	0.63	0.14	0.01	0.16	0.43
ORT	0.43	0.46	0.40	0.88	0.56	0.63	0.88	0.72
WIT	0.08	0.77	0.37	0.02	0.69	0.27	0.04	0.46

Notes: GMM estimates of the log-linear forward-looking rational-addiction demand model. Dependent variable: log legal quantity sold. Endogenous regressors: future legal quantity and the contemporaneous and future own price. Continuous regressors are in logs and, where applicable, real terms. All specifications include monthly seasonal dummies, department fixed effects, and department-specific year effects. Wild cluster bootstrap standard errors, clustered by department, in parentheses. The Total column refers to aggregate legal sales in the observed segment; its price variable is the sales-weighted average final price. SER: standard error of the regression; DW: Durbin–Watson statistic; DWH and ORT: p-values of the Durbin–Wu–Hausman endogeneity test and the Hansen J overidentification test. WIT: the Lewis and Mertens (2025) weak-instrument statistic, increasing in instrument strength—values above the 5% critical value reject weak identification, values below indicate weak instruments. The critical value varies with the number of endogenous regressors and instruments but exceeds 20 throughout; all WIT values here fall well below it, so the test fails to reject weak identification in every column.

Table 3 reports the estimates. The results should be interpreted with caution. The forward-looking specification is informative as a diagnostic exercise, but the available instruments do not identify the model with enough precision to make the implied elasticities a reliable basis for policy evaluation. Several coefficients are imprecisely estimated, and the results are less stable than in the static and lagged-demand specifications. This is not surprising. A forward-looking rational-addiction model asks the data to identify more objects—persistence, anticipation, contemporaneous price effects, and future-price effects—using a relatively short monthly panel and a limited set of instruments.

The main lesson from Table 3 is therefore negative but useful. The rational-addiction framework is conceptually relevant, but the available data do not support treating the forward-looking specification as the preferred empirical model. The static and lagged-demand specifications provide clearer and more stable evidence on legal demand in the observed market. The forward-looking estimates do not overturn the central result: brand-level legal demand is much more price-responsive than aggregate legal demand, and aggregate elasticities are a poor guide to the substitution margins that determine the vulnerability of the taxed legal base.

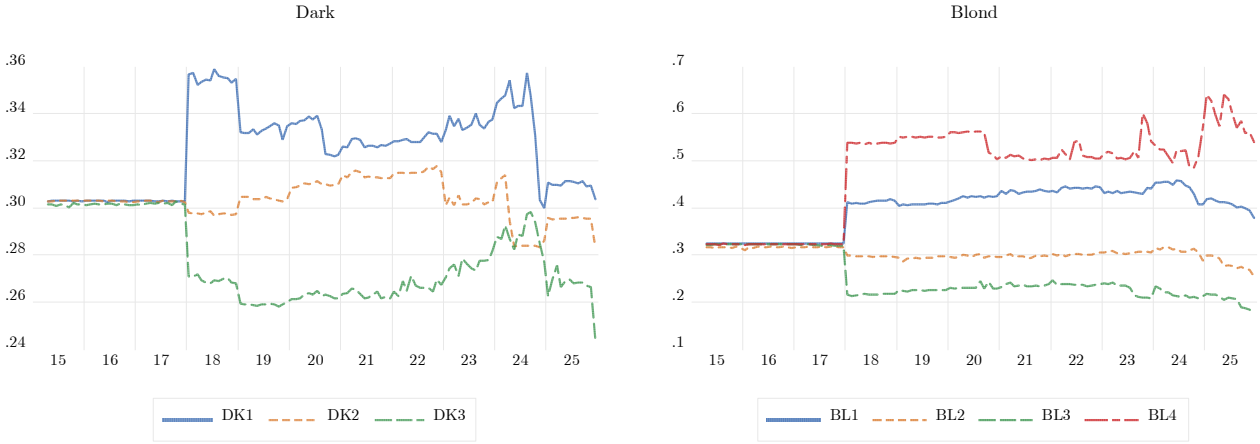
The interpretation is also limited by the outcome being modeled. Equation (16) describes legal tax-paid purchases in the observed segment. Even if the forward-looking model were precisely identified, it would still estimate dynamic responses of legal demand, not total cigarette consumption.

5.3 Effective tax burdens and price adjustment

The 2018 reform changed the way cigarette excises affected legal brands. Under the previous *ad valorem* regime, tax liability was tied to value. Under the *ad quantum* regime, tax liability became tied to physical units, with different amounts for dark and blond cigarettes. This distinction matters because, within each tax category, a unit tax represents a larger effective tax share for lower-price brands. The reform therefore changed relative effective tax burdens across legal brands before considering any subsequent adjustment by firms or consumers.

Figure 2

Effective ICE burden by brand: ICE payments as a share of gross revenue



Notes: Monthly national series of the ratio of ICE payments to gross revenue for each brand. The left panel corresponds to dark cigarette brands and the right panel to blond cigarette brands. Departmental ICE payments and gross revenue are aggregated within each brand. The vertical scales differ across panels.

Figure 2 reports the effective ICE burden by brand, measured as ICE payments relative to gross revenue. Before the reform, effective ICE burdens were comparatively stable within each tobacco category, consistent with the proportional nature of the *ad valorem* regime. After the reform, the burden became more uneven. The pattern is especially pronounced among blond cigarettes, where the effective ICE burden displays larger differences across brands after the adoption of category-specific unit taxation. Among dark cigarettes, the vertical scale is narrower, but the same principle applies: the unit tax altered effective burdens within the category.

This measure should not be interpreted as economic incidence in the full sense. It does not identify how the burden is ultimately divided among consumers, firms, suppliers, or other agents. It is an accounting measure of the tax share embedded in observed legal transactions. For this paper, however, that object is directly relevant. Excise revenue is collected from legal transactions, and the effective tax share helps characterize how the tax regime affected different legal products.

Figure 3

Real gross and net-of-ICE price indexes by brand (April 2015=100)

Notes: Monthly national indexes of real gross prices and real prices net of ICE by brand, with April 2015 = 100. Gross prices are constructed as gross revenue divided by quantity sold. Net-of-ICE prices are constructed as net revenue divided by quantity sold. Both are deflated using the national consumer price index. The top panels correspond to real gross prices and the bottom panels to real net-of-ICE prices. The left panels correspond to dark cigarette brands and the right panels to blond cigarette brands. Departmental revenues and quantities are aggregated within each brand before constructing the monthly unit values. The vertical scales differ across panels.

Figure 3 complements this evidence by showing real gross and net-of-ICE price indexes by brand. Gross prices capture the final prices paid in the legal market. Net-of-ICE prices capture revenue per unit before the excise payment. Comparing both series is useful because tax changes need not appear only as uniform increases in final consumer prices. In a differentiated market, firms may adjust net-of-tax prices differently across brands when demand conditions, competitive pressure, and substitution possibilities differ.

Table 4

Effective tax burdens and real price adjustment

Variable	DK1	DK2	DK3	BL1	BL2	BL3	BL4	Total
ICE burden	2.96 (0.87)	0.08 (0.79)	-3.26 (0.62)	10.14 (1.97)	-1.86 (1.37)	-9.83 (1.66)	21.38 (1.12)	0.35 (0.97)
Gross Price	0.95 (5.38)	-6.58 (6.80)	5.18 (6.10)	-3.50 (5.49)	3.12 (3.06)	2.69 (1.79)	1.45 (10.4)	0.95 (5.34)
Net Price	-3.65 (2.49)	-6.76 (5.75)	9.49 (7.31)	-20.03 (4.55)	21.40 (1.37)	41.29 (1.82)	-36.72 (8.34)	8.37 (6.93)

Notes: Brand-level changes after the reform in the effective ICE burden and in real gross and net-of-ICE prices. The effective ICE burden is measured as ICE payments divided by gross revenue; changes in this row are expressed in percentage points. Gross and net-of-ICE price entries are coefficients from regressions of log real prices on a constant and a post-2017 dummy, multiplied by 100. These coefficients are log-point differences between the pre- and post-reform periods, not exact percentage changes. They indicate the direction and statistical significance of average price adjustment; for small values they approximate percentage changes. HAC standard errors are reported in parentheses.

Table 4 summarizes this adjustment numerically. For each brand, it reports the change in the effective ICE burden and the estimated post-reform change in real gross and net-of-ICE prices. The price entries are obtained from regressions of log real prices on a constant and a post-2017 dummy, with HAC standard errors. They should be interpreted as log-point differences, not exact percentage changes. The exercise is intentionally simple: it summarizes the average price adjustment shown in Figure 3; it does not estimate a structural pass-through coefficient.

The results point to a clear pattern at the extremes of the price ladder. Lower-price brands within their segment—DK1, BL1, and BL4—experienced increases in the effective ICE burden and negative net-of-ICE price log differences. Higher-price brands—especially DK3 and BL3, and to some extent BL2—experienced reductions in the effective ICE burden and positive net-of-ICE price log differences. By contrast, the gross-price log differences are not statistically significant. The strongest adjustment is therefore observed in net-of-tax prices rather than in final consumer prices.

The absence of statistically significant changes in gross real prices should be interpreted with care. It does not imply that the reform was irrelevant for the legal market. The reform appears to have changed the structure of effective taxation more than the average

tax burden of the legal portfolio. Its main effect was to redistribute tax pressure across brands: lower-price legal brands faced higher effective ICE burdens, while some higher-price brands faced lower effective burdens. This redistribution is more visible in net-of-tax prices than in gross prices.

This pattern is economically informative. Lower-price legal brands are more likely to compete directly with untaxed or illicit alternatives. That competition limits the ability to raise gross consumer prices when the effective tax burden increases. If gross prices do not rise significantly, a higher tax share must be absorbed through lower net-of-tax prices. Higher-price brands face a different competitive environment. Their lower effective burden after the reform, combined with weaker direct exposure to low-price untaxed alternatives, is consistent with stronger net-price trajectories. The evidence therefore suggests heterogeneous margin adjustment within the legal brand portfolio: margin compression in lower-price brands and stronger margins in higher-price brands.

At the same time, legal quantities fell sharply. This is the key tension. The adjustment was not a simple story of higher average consumer prices reducing legal demand. Gross real prices do not display statistically clear average increases, yet the legal tax-paid base contracted substantially. The evidence is more consistent with a broader reallocation in which relative effective burdens, net margins, brand composition, and substitution possibilities all mattered.

This has direct implications for the fiscal claim. The reform did not operate as a simple aggregate tax shock that uniformly raised consumer prices, reduced smoking, and increased revenue. It mainly changed relative effective burdens within the legal market. In a differentiated market with brand substitution and untaxed alternatives, such a reform may erode the legal base even without a clear sustained increase in average real consumer prices. The observed fall in legal quantities and real ICE revenue is therefore central: the taxed base contracted, while the mechanism needed to infer a proportional fall in total smoking remains unobserved.

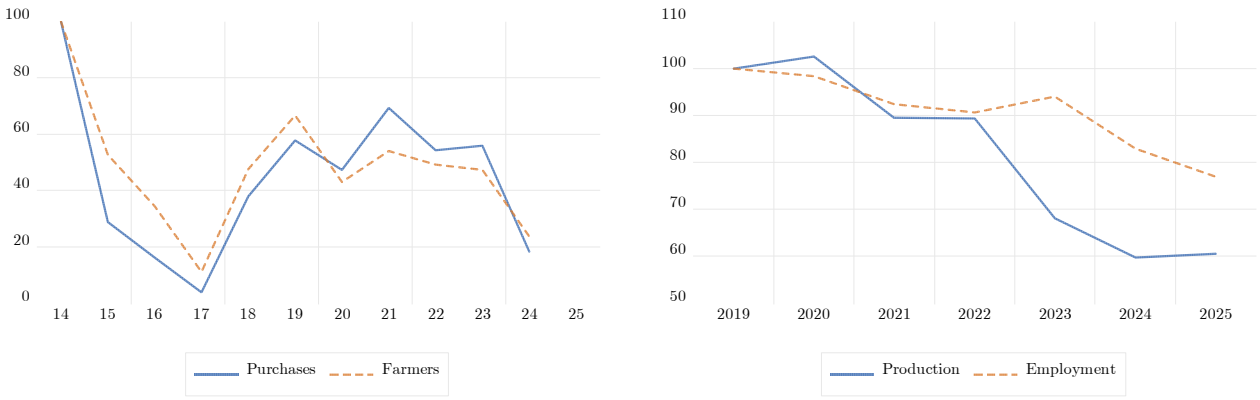
Figures 2 and 3, together with Table 4, therefore provide the empirical counterpart to the pricing block in Section 3.2. Statutory tax liabilities, effective tax shares, gross prices, net-of-tax prices, and legal quantities are distinct objects. Revenue depends on the legal base that remains after firms and consumers adjust, not on statutory tax rates alone. The evidence does not estimate structural pass-through, but it shows why statutory tax arithmetic is insufficient for evaluating the fiscal performance of the reform.

5.4 The legal production chain

The contraction of the observed legal market was also reflected in the formal production chain. Figure 4 reports annual indicators for production, employment, labor costs, and purchases of dark tobacco from local farmers. These variables are not used to estimate demand. They are used to assess whether the weakening of the taxed legal segment was confined to final sales and tax payments or whether it was also present upstream and downstream in the formal sector.

Figure 4

Indicators of the legal production chain



Notes: Indexes normalized to 100 in the first year for which information is available for each series. The left panel corresponds to upstream indicators: purchases of dark tobacco and the number of farmer families from which tobacco was purchased. The right panel corresponds to downstream indicators: cigarette production and employment at CITSA for the brands it produces and sells. The vertical scales differ across panels.

The figure shows a clear deterioration in formal-chain indicators over the sample period. Production declined, employment fell, labor costs weakened, and purchases of dark tobacco

from domestic farmers contracted. These patterns are consistent with the broader contraction of the legal tax-paid segment documented above. The point is not that each of these changes can be attributed mechanically to the tax reform alone. Nor is it that a decline in formal tobacco-sector activity is automatically a welfare loss. In a general-equilibrium setting, resources displaced from one sector may be reallocated elsewhere.

The relevance of Figure 4 is different. It shows that the erosion of the legal cigarette market was not only a consumer-side phenomenon. The contraction was also visible in the formal structure through which legal cigarettes are produced, inputs are purchased, workers are employed, and taxes are collected. This matters for the fiscal claim because the fiscal promise of excise taxation depends on the resilience of the legal tax base. If higher excises are associated with lower tax-paid sales, lower real revenue, and weaker formal-chain indicators, the adjustment is broader than a simple reduction in smoking.

This evidence should therefore be read as part of the fiscal and institutional evaluation of the reform. It does not identify total cigarette consumption or the size of the illicit market. It does, however, reinforce the distinction between total smoking and the taxed legal base. A shrinking legal base can weaken formal activity even when the effect on total consumption remains unobserved. For that reason, the legal production chain is relevant not because it delivers a stand-alone welfare conclusion, but because it shows how legal-base erosion can propagate beyond final sales and tax payments.

5.5 The taxed base and the "magic bullet"

The evidence presented so far identifies a contraction of the observed legal market. Legal sales declined, real ICE revenue fell, brand-level legal demand is much more price-responsive than aggregate legal demand, effective tax burdens and prices adjusted heterogeneously, and the legal production chain weakened. These facts are directly relevant for fiscal evaluation because excise revenue is collected on legal tax-paid sales. They do not, however, identify the response of total cigarette consumption.

What remains open is not the behavior of the tax base, but the behavior of total cigarette consumption. Section 3.3 already established why these two objects need not move together. Since $Q_t^L = \theta_t Q_t$, it follows that

$$dQ_t^L = \theta_t dQ_t + Q_t d\theta_t,$$

or equivalently:

$$dQ_t = \frac{1}{\theta_t}\Big[dQ_t^L - Q_t d\theta_t\Big]. \tag{18}$$

Thus, the observed decline in legal sales coincides with the decline in total consumption only if the legal share does not fall. That is a strong assumption in the present setting. A fall in legal sales may reflect a fall in total cigarette consumption, a fall in the legal share, or both. If the legal share falls sufficiently, total consumption may fall by less than legal sales, remain roughly unchanged, or even increase.

The empirical evidence suggests that the legal-share margin cannot be ignored. First, brand-level legal demand is much more price-responsive than aggregate legal demand. This indicates that consumers reallocate across products in ways that aggregate data conceal. Second, the broader cigarette-price proxy and the Paraguayan price proxy generally enter with positive coefficients in the brand-level equations, indicating that demand for a given legal brand tends to rise when competing cigarette alternatives become more expensive. This is consistent with substitution between legal cigarettes and external, potentially untaxed alternatives.[3]

This evidence identifies substitution from the legal-demand side, but it does not identify the reverse effect from legal prices to illicit demand. Marshallian cross-price effects need not be symmetric. A positive response of legal demand to a higher external cigarette price does not mechanically prove that illicit demand rises when legal prices rise. Still, a sign

[3] The exception is BL4, whose Paraguayan-price coefficient is negative (-0.63 and -0.86 in the static and lagged-demand specifications). As the lowest-price blond brand, BL4 is the legal product most exposed to low-price Paraguayan-origin cigarettes, for which a single representative Paraguayan price is a poor proxy; the negative sign is better read as proxy limitation than as genuine complementarity. It does not affect the legal-share argument, which rests on the general pattern across brands.

reversal would require strong and asymmetric income effects. In particular, illicit cigarettes would have to be sufficiently inferior relative to legal brands for a rise in legal prices to reduce illicit consumption through the income effect. The estimated income effects in Tables 1 and 2 are limited and heterogeneous across brands, which provides little support for such an interpretation.

The more plausible working interpretation is that legal and untaxed cigarettes are substitutes, especially for price-sensitive consumers facing changes in relative prices. Under that interpretation, an increase in the relative attractiveness of untaxed products reduces the legal share of consumption. Total cigarette consumption then falls less than legal sales. If the decline in the legal share is large enough, total consumption may remain roughly unchanged or even increase. This remains a maintained interpretation, not a directly identified result. Its importance is that it rejects the mechanical equivalence between lower legal sales and lower total smoking.

The conclusion is therefore deliberately narrower than the public-health interpretation often attached to excise increases. The paper identifies erosion of the taxed legal base. It does not identify an equivalent reduction in total cigarette consumption. The fiscal component of the strong "magic bullet" claim fails in the observed legal market because real excise revenue fell. The health component cannot be inferred from legal sales alone. Turning these observed legal-market outcomes into a welfare statement requires additional assumptions, which are discussed next.

5.6 Welfare implications

The previous section showed that the empirical evidence identifies erosion of the taxed legal base. This distinction is central for welfare. A contraction in legal sales implies a loss of private surplus in the observed legal market, unless consumers preserve surplus by substituting to other products or channels. From a social-welfare perspective, such a contraction can be justified only if the externality or internality gains associated with lower total smoking are large enough to offset private surplus losses, fiscal losses, enforcement costs, and any additional costs generated by displacement toward untaxed

products. Welfare conclusions therefore require assumptions beyond the observed legal-market data.

The accounting framework in Section 3.4 approximates the welfare change as in equation (11):

$$\Delta W \approx \Delta CS^{L} + \Delta CS^{I} + \eta \Delta R - \nu \Delta Q - \kappa \Delta Q^{I} - \Delta e.$$

Some components of this equation can be measured or approximated with the available data, while others are unobserved. The legal-market component of consumer surplus, ΔCS^{L}, can be approximated using observed legal prices and quantities. The change in real excise revenue, ΔR, is observed in the legal market. By contrast, the surplus change associated with illicit or untaxed consumption, ΔCS^{I}, is not observed. Nor are the per-unit uninternalized harm ν, the social cost of illicit-market expansion κ, the change in enforcement costs Δe, or the social value of public funds η.

The legal-market component of consumer surplus is approximated with a trapezoidal compensating-variation calculation. For each brand n, let $\overline{p}_{n,pre}$ and $\overline{p}_{n,post}$ denote average real gross prices before and after the reform, and let $\overline{q}_{n,pre}$ and $\overline{q}_{n,post}$ denote the corresponding average monthly legal quantities. The legal-market compensating variation is approximated by

$$\widehat{CV}^{L} \approx \frac{1}{2} \sum_{n} \left(\overline{p}_{n,post} - \overline{p}_{n,pre} \right) \left(\overline{q}_{n,pre} + \overline{q}_{n,post} \right),$$

so that

$$\widehat{\Delta CS}^{L} = -\widehat{CV}^{L}.$$

The calculation uses gross prices because those are the prices faced by consumers. It is aggregated across brands rather than computed from the consolidated unit value, because a central empirical result of the paper is that aggregation masks brand-level reallocation. The resulting object is not total consumer surplus. It measures only the surplus change associated with the observed legal segment.

To make the break-even logic operational, consider first a benchmark that is deliberately favorable to the reform. Set $\Delta CS^I = \kappa = \Delta e = 0$. These assumptions ignore any loss of surplus from consumers who move to untaxed products, any resource or social costs associated with illicit-market expansion, and any additional enforcement or administrative costs. Set also $\eta = 1$, so that a unit of public revenue has the same social value as a unit of private surplus. Finally, assume that the entire decline in legal sales corresponds to a decline in total cigarette consumption, so that $\Delta Q = \Delta Q^L < 0$. This is the most favorable consumption assumption because it attributes all missing legal cigarettes to lower total smoking rather than to substitution toward untaxed products.

Under these assumptions, the break-even value in equation (12) collapses to the following legal-market break-even statistic:

$$\nu^* = \frac{\widehat{\Delta CS}^L + \Delta R}{\Delta Q^L} \equiv b. \tag{19}$$

Equation (19) measures the uninternalized harm per cigarette that would be required to offset the observed legal-market losses under the favorable benchmark. To operationalize it, $\widehat{\Delta CS}^L$, ΔR, and ΔQ^L are computed as pre/post differences using average monthly values.[4] Thus, ΔR is the change in average monthly real excise revenue, and ΔQ^L is the change in average monthly legal quantities.

Like ΔR and ΔQ^L, the compensating variation is a difference between pre- and post-reform period averages, so it embeds the broader downward trend in legal quantities and not only a reform-induced movement along legal demand. The figure is thus the surplus change over the legal segment between the two periods, and b a period-comparison break-even threshold, not a causal decomposition of the reform's welfare effect. This is consistent with the accounting interpretation maintained throughout the section.

Rather than reporting b in currency units, it is more useful to express it relative to the average gross legal price per stick in December 2025, denoted by $\overline{p}_{2015:12}$. Define

[4] The pre-reform period is 2015–2017, and the post-reform period is 2018–2025.

$$\tilde{b} = \frac{b}{\overline{p}_{2015:12}}.$$

Using brand-level calculations of (19) and the average price per stick, $\tilde{b} = 0.3..$ Thus, under the favorable benchmark, the uninternalized harm required to offset the observed legal-market losses is about 30 percent of the December 2025 average legal price per stick. This threshold coincides with the pre-reform effective ICE burden.

The favorable benchmark can then be relaxed by allowing only part of the contraction in legal sales to represent lower total smoking. Define

$$\psi = \frac{-\Delta Q}{-\Delta Q^{L}},$$

where ψ is the fraction of the decline in legal sales that corresponds to a decline in total cigarette consumption. If $\psi = 1,$ every missing legal cigarette is one less cigarette consumed. If $0 < \psi < 1,$ part of the legal contraction reflects substitution toward untaxed products. If $\psi = 0,$ total consumption does not fall. If $\psi < 0,$ total consumption increases even though legal sales fall.

Under the same benchmark normalization for the other unobserved terms, the break-even value becomes:

$$\frac{\nu^{*}\left(\psi\right)}{\overline{p}_{2015:12}} = \frac{\tilde{b}}{\psi} \simeq \frac{0.3}{\psi}.$$

If all missing legal cigarettes represent lower total consumption $\left(\psi = 1\right),$ the break-even uninternalized harm is about 30 percent of the average legal price. . Roughly one third of this threshold reflects the deterioration of legal consumer surplus, while the remaining two thirds reflect lost fiscal revenue. If only half of the decline in legal sales reflects lower total consumption $\left(\psi = 0.5\right),$ the required harm doubles to about 60 percent of the average price. If only one quarter does $\left(\psi = 0.25\right),$ the threshold rises to about 120 percent of the average price. If $\left(\psi = 0\right),$ the reduced-smoking channel cannot justify the reform with any externality.

These figures are not welfare estimates. They are break-even values under explicit assumptions. Their usefulness is that they show how demanding the welfare case becomes as the legal-share margin becomes more important. If ν is interpreted narrowly as an externality, the benchmark threshold is already high, and it becomes much harder to defend the reform as ψ falls. A welfare defense becomes more plausible only if total consumption fell substantially or if ν includes sizable internalities. But the latter requires additional behavioral and normative assumptions that are not identified from the legal-market data.

The conclusion is deliberately limited. The observable legal-market components point against a welfare gain: legal consumer surplus falls and real excise revenue falls. A social-welfare gain must therefore come from unobserved components, especially externality or internality benefits associated with lower total smoking. But total smoking is not observed, and neither are illicit consumer surplus, illicit-market costs, enforcement costs, or the social value of public funds. The calculation obtained is favorable to the reform in several dimensions: it assumes no loss of surplus from substitution toward untaxed products, no social cost of illicit-market expansion, and no additional enforcement or administrative costs. In a setting with high informality, porous borders, weak enforcement capacity, and available evidence of illicit-market penetration, these assumptions are unlikely to be innocuous. If only part of the fall in legal sales reflects lower total smoking, the required uninternalized harm rises quickly; if total smoking does not fall, the externality and internality channels do not operate. The accounting exercise therefore does not establish the welfare effect of the reform. It shows that any welfare defense must make explicit assumptions about total consumption, illicit substitution, private surplus, uninternalized harms, enforcement costs, illicit-market costs, and the value of public funds. Automatic equivalences between lower legal sales, lower smoking, and higher welfare are not valid.

6 Concluding remarks

Tobacco excises are often presented as a policy instrument capable of achieving several objectives at once: raising consumer prices, reducing smoking, increasing fiscal revenue, and improving welfare. This paper argues that this "magic bullet" interpretation is too

strong unless the relevant margins are separated. A statutory tax increase is only the first step. Its effects depend on how taxes alter effective tax burdens and prices, how consumers substitute across legal brands and toward untaxed alternatives, how the taxed legal base responds, and whether any reduction in total smoking generates welfare gains large enough to offset private surplus losses, fiscal losses, enforcement costs, and other adjustment costs.

The Bolivian experience illustrates why these distinctions matter. The 2018 reform shifted cigarette excise taxation from an *ad valorem* regime to a category-specific *ad quantum* regime. This changed the structure of effective tax burdens within the legal market. Lower-price brands within their category faced higher effective ICE burdens, while some higher-price brands faced lower effective burdens. The adjustment was not reflected in statistically clear increases in gross real prices. It appeared more strongly in net-of-tax prices: lower-price brands experienced margin compression, while higher-price brands displayed stronger net-price trajectories. This is inconsistent with a simple mechanism in which higher excises are uniformly transmitted into final consumer prices.

The demand estimates reinforce this interpretation. Brand-level legal demand is generally much more price-responsive than aggregate legal demand. The aggregate elasticity is small, but this is not the policy-relevant object if aggregation hides substitution across brands, segments, and channels. Aggregate unit values combine within-brand price changes with shifts in brand composition and may conceal precisely the margins that determine the vulnerability of the taxed legal base. This also cautions against interpreting elasticities from household expenditure surveys mechanically when "prices" are constructed as expenditure divided by quantity and when illicit purchases or brand composition are not properly observed.

The fiscal result is direct. In the legal segment covered by the data, legal sales fell and real ICE revenue fell. Thus, the fiscal component of the strong "magic bullet" claim did not materialize in the observed legal market. The contraction of the taxed legal segment was also accompanied by deterioration in the formal production chain, including production, employment, labor costs, and purchases from domestic suppliers. These outcomes are not welfare losses by themselves, because resources may be reallocated in

general equilibrium. They do show, however, that the adjustment was broader than final sales and tax payments.

The health result is not identified from the same evidence. The data observe legal tax-paid sales, not total cigarette consumption. If the legal share of consumption remained constant, the decline in legal sales would correspond to a proportional decline in total smoking. If the legal share fell, total consumption declined by less than legal sales. If the legal share fell sufficiently, total consumption may have remained unchanged or even increased. The evidence is consistent with substitution across brands and possibly toward untaxed alternatives, but it does not identify illicit quantities. The paper therefore identifies erosion of the taxed legal base, not an equivalent reduction in total smoking.

The welfare calculation strengthens this caution. The observable legal-market components move against the reform: legal consumer surplus falls and real excise revenue falls. Under the most favorable consumption assumption, in which all missing legal cigarettes correspond to lower total consumption, the break-even uninternalized harm is about 30 percent of the average legal price. Roughly one third of this threshold reflects the deterioration of legal consumer surplus, while the remaining two thirds reflects lost fiscal revenue under the normalization that public and private funds have equal social value. This benchmark is already favorable to the reform. It assumes no loss of surplus from substitution toward untaxed products, no social cost from illicit-market expansion, and no additional enforcement or administrative costs. In Bolivia's institutional setting—high informality, porous borders, weak enforcement capacity, and available evidence of illicit-market penetration—these assumptions are unlikely to be innocuous. If only part of the decline in legal sales reflects lower total smoking, the required uninternalized harm rises quickly; if total smoking does not fall, neither the externality nor the internality channel operates. A welfare defense therefore requires strong and explicit assumptions about total consumption, illicit substitution, uninternalized harms, enforcement costs, illicit-market costs, and the value of public funds. It cannot be inferred from the fall in legal sales alone.

The broader lesson is that excise policy should be evaluated through the tax base that is actually taxed, the prices consumers actually face, and the substitution margins available

to consumers and firms. Higher statutory tax rates do not mechanically imply higher revenue, lower total consumption, or higher welfare. In markets where legal and untaxed products coexist, and where consumers can move across brands and channels, excise reforms may erode the legal base precisely in the segments where the effective burden rises most. That does not imply that tobacco taxation cannot be justified. It implies that the justification must be empirical and explicit: it must show how much total smoking falls, how much of the legal base is preserved, and how welfare gains compare with the private and fiscal costs generated by the reform.

References

Arteaga, S., M. Chumacero, I. Finot, C. Garín, J. Morales, G. Paraje, L. Pruzzo, and M. Vargas (2024). "Caracterización de los Hábitos de Consumo de Tabaco en Bolivia," *Manuscript*, SDSN Bolivia and Universidad Privada Boliviana.

Badenes-Plá, N. and A. Jones (2003). "Addictive Goods and Taxes: A Survey from an Economic Perspective," *Hacienda Pública Española / Revista de Economía Pública*, 167(4), 123–151.

Becker, G., M. Grossman, and K. Murphy (1994). "An Empirical Analysis of Cigarette Addiction," *The American Economic Review*, 84(3), 396–418.

Becker, G. and K. Murphy (2006). "The Market for Illegal Goods: The Case of Drugs," Journal of Political Economy, 114(1), 38–60.

Cameron, A., J. Gelbach, and D. Miller (2008). "Bootstrap-Based Improvements for Inference with Clustered Errors," *The Review of Economics and Statistics*, 90(3), 414–427.

Cameron, A. and D. Miller (2015). "A Practitioner's Guide to Cluster-Robust Inference," *Journal of Human Resources*, 50(2), 317–372.

Chaloupka, F. and K. Warner (2000). "The Economics of Smoking," in A. Culyer and J. Newhouse (eds.), *Handbook of Health Economics*, Vol. 1B. Elsevier.

Chetty, R. (2009). "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 1, 451–487.

Cicowiez, M., G. Cruces, G. Falcone, and J. Puig (2024). "On the Impacts of Higher Tobacco Taxes in Argentina: A Computable General Equilibrium Approach," *Tobacco Control*, 33, s115–s121.

Cotti, C., C. Courtemanche, J. Maclean, E. Nesson, M. Pesko, and N. Tefft (2022). "The Effects of E-cigarette Taxes on E-cigarette Prices and Tobacco Product Sales: Evidence from Retail Panel Data," *Journal of Health Economics*, 86, 102676.

Dare, C., M. Boachie, E. Tingum, S. Abdullah, and C. van Walbeek (2021). "Estimating the Price Elasticity of Demand for Cigarettes in South Africa Using the Deaton Approach," *BMJ Open*, 11:e046279.

Davidson, R. and J. MacKinnon (2010). "Wild Bootstrap Tests for IV Regression," *Journal of Business & Economic Statistics*, 28(1), 128–144.

Deaton, A. (1988). "Quality, Quantity, and Spatial Variation of Price," *The American Economic Review*, 78(3), 418-430.

DeCicca, P., D. Kenkel, and F. Liu (2013). "Excise Tax Avoidance: The Case of State Cigarette Taxes," *Journal of Health Economics*, 32, 1130-1141.

DeCicca, P., D. Kenkel, and M. Lovenheim (2022). "The Economics of Tobacco Regulation: A Comprehensive Review," *Journal of Economic Literature*, 60(3), 883–970.

Delipalla, S. and O. O'Donnell (2001). "Estimating Tax Incidence, Market Power and Market Conduct: The European Cigarette Industry," *International Journal of Industrial Organization*, 19, 885–908.

Farhi, E. and X. Gabaix (2020). "Optimal Taxation with Behavioral Agents," *American Economic Review*, 110(1), 298–336.

Ferre, Z., M. Gerstenblüth, and P. Triunfo (2025). "Demand for Cigarettes in Uruguay: a National Analysis with Monthly Data," *Discover Public Health*, 22: 98, https://doi.org/10.1186/s12982-025-00495-9.

Gibson, J. and B. Kim (2019). "The Price Elasticity of Quantity, and of Quality, for Tobacco Products," *Health Economics*, 28, 587–593.

Gilleskie, D. (2014). "Dynamic Models: Econometric Considerations of Time," in A. Culyer (ed.), *Encyclopedia of Health Economics*, 1, Elsevier.

Gonçalves, S., and L. Kilian (2004). "Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form," *Journal of Econometrics*, 123(1), 89–120.

Gruber, J. and B. Kőszegi (2004). "Tax Incidence When Individuals Are Time-Inconsistent: The Case of Cigarette Excise Taxes," *Journal of Public Economics*, 88, 1959–1987.

Gruber, J., A. Sen, and M. Stabile (2003). "Estimating Price Elasticities when There is Smuggling: The Sensitivity of Smoking to Price in Canada," *Journal of Health Economics*, 22(5), 821–842.

Hendry, D. (1995). *Dynamic Econometrics*, Oxford University Press.

ILOSTAT (2024). *Country profiles: Bolivia*. International Labour Organization. https://ilostat.ilo.org/data/country-profiles/bol/

Jha, P. and F. Chaloupka (editors) (2000). *Tobacco Control in Developing Countries*, Oxford University Press.

John, R., V. Vulovic, G. Chelwa, and F. Chaloupka (2023). "Updated Toolkit on Using Household Expenditure Surveys for Research in the Economics of Tobacco Control. A Tobacconomics Toolkit". Tobacconomics, Institute for Health Research and Policy, University of Illinois Chicago.

Joossens, L. and M. Raw (2008). "Progress in Combating Cigarette Smuggling: Controlling the Supply Chain," *Tobacco Control*, 17, 399–404.

Jürges, H. (2004). "The Welfare Costs of Addiction," *Journal of Contextual Economics – Schmollers Jahrbuch*, 124(3), 327-353.

Lewis, D. and K. Mertens (2025). "A Robust Test for Weak Instruments for 2SLS with Multiple Endogenous Regressors," *The Review of Economic Studies*, doi:10.1093/restud/rdaf103.

Lovenheim, M. (2008). "How Far to the Border?: The Extent and Impact of Cross–Border Casual Cigarette Smuggling Revenues," *National Tax Journal*, 61(1), 7–33.

MacKinnon, J., M. Nielsen, and M. Webb (2023). "Cluster-Robust Inference: A Guide to Empirical Practice," *Journal of Econometrics*, 232(2), 272–299.

Mammen, E. (1993). "Bootstrap and Wild Bootstrap for High Dimensional Linear Models," *The Annals of Statistics*, 21(1), 255–285.

Morales, J. and S. Santander (2025). "Assessing Cigarette Reduction Tax-Effectiveness in Low Tobacco Expenditure Contexts: An Application to Bolivia," *Health Economics*, 34(4), 758-779.

O'Donoghue, T. and M. Rabin (2006). "Optimal Sin Taxes," *Journal of Public Economics*, 90, 1825–1849.

OECD. (2024). *Tobacco Taxation in Latin America and the Caribbean*, OECD Publishing.

Orphanides, A. and D. Zervos (1995). "Rational Addiction with Learning and Regret," *Journal of Political Economy*, 103(4), 739-758.

Paraje, G, D. Araya, A. De Paz, and N. Nargis (2021). "Price and Expenditure Elasticity of Cigarette Demand in El Salvador: A Household-level Analysis and Simulation of a Tax Increase," *Tobacco Control*, 30, 422–427.

Saenz de Miera Juarez, B., L. Reynales-Shigematsu, M. Stoklosa, K. Welding, and J. Drope (2021). "Measuring the Illicit Cigarette Market in Mexico: A Cross Validation of Two Methodologies," *Tobacco Control*, 30(2), 125–131.

Slemrod, J. and S. Yitzhaki (2002). "Tax Avoidance, Evasion, and Administration," in A. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics*, Vol. 3, Chapter 22, 1423–1470. Elsevier.

Tauras, J., M. Pesko, J. Huang, F. Chaloupka, and M. Farrelly (2016). "The Effect of Cigarette Prices on Cigarette Sales: Exploring Heterogeneity in Price Elasticities at High and Low Prices," *NBER Working Paper* No. 22251.

Appendix A
Endogeneity and instrumental-variable diagnostics

A natural concern in the estimation of legal cigarette demand is that the brand's own price, $p_{n,d,t}$, may be endogenous. Prices and quantities are equilibrium outcomes, so demand shocks may induce price adjustments. To evaluate this concern, equation (13) is re-estimated by instrumental variables, treating the brand's own price as endogenous. The excluded instruments are the real price of electricity and the real average wage in the tobacco industry, both obtained from INE. These variables are interpreted as supply-side cost shifters that may affect legal cigarette prices without directly entering the demand equation, conditional on the remaining controls and fixed effects.

Table A.1 reports the IV estimates. The Durbin–Wu–Hausman tests do not provide systematic evidence against treating own prices as exogenous in the conditional demand equations: the null is rejected only for BL3 and BL4. The Hansen overidentifying-restriction tests also generally fail to reject the validity of the overidentifying restrictions, with the main exception of DK2. Thus, the diagnostic tests do not provide broad evidence against the conditional specification.

The main limitation of the IV exercise is weak identification. The Lewis and Mertens (2025) weak-instrument statistic fails to reject weak identification in all specifications except DK3. This is consistent with the pattern of estimates: several IV own-price coefficients are imprecisely estimated and some have implausible positive signs. As a result, the IV estimates do not provide a stable or better-identified alternative to the static conditional benchmark reported in Table 1.

A secondary concern is the broader cigarette-category price proxy, $P_{d,t}$. Specifications treating both $p_{n,d,t}$ and $P_{d,t}$ as endogenous were also examined using the same excluded instruments. These estimates were more unstable and did not materially change the conclusions, so they are not reported.

Table A.1

IV estimates of legal cigarette demand by brand and in the aggregate

Variable	DK1	DK2	DK3	BL1	BL2	BL3	BL4	Total
$p_{n,d,t}$	-2.20 (4.76)	1.32 (5.69)	-1.17 (0.91)	-3.37 (3.26)	-0.88 (0.57)	4.48 (2.59)	1.36 (1.15)	-0.39 (0.57)
P_t	0.68 (0.37)	-0.59 (0.76)	0.39 (0.12)	0.53 (0.39)	0.30 (0.09)	-0.18 (0.27)	0.71 (0.36)	0.27 (0.07)
p_t^I	0.46 (0.57)	2.78 (0.73)	0.37 (0.22)	0.74 (0.47)	0.28 (0.21)	0.32 (0.40)	-0.62 (0.34)	0.17 (0.15)
$p_{d,t}^{coffee}$	0.27 (0.52)	0.18 (0.27)	0.11 (0.08)	-0.10 (0.13)	-0.05 (0.11)	-0.87 (0.33)	0.13 (0.24)	0.00 (0.10)
$p_{d,t}^{beer}$	0.66 (0.55)	-0.16 (0.76)	0.49 (0.25)	-0.39 (0.41)	-0.20 (0.20)	-2.35 (1.03)	0.55 (0.90)	0.00 (0.17)
$p_{d,t}^{coca}$	-0.28 (0.14)	-0.47 (0.48)	-0.20 (0.06)	-0.04 (0.03)	-0.01 (0.07)	-0.12 (0.16)	-0.43 (0.12)	-0.05 (0.05)
$p_{d,t}^{gum}$	-0.13 (1.10)	0.28 (0.43)	-0.06 (0.11)	0.29 (0.51)	-0.03 (0.06)	-0.01 (0.45)	0.76 (0.26)	-0.06 (0.05)
y_t	1.47 (2.18)	-0.43 (5.03)	-1.09 (0.34)	1.88 (1.73)	0.45 (0.52)	-1.54 (1.49)	-3.99 (1.31)	-0.28 (0.33)
Observations	1161	1107	1161	1161	1161	1156	1161	1161
Adjusted R^2	0.98	0.94	0.98	0.99	0.98	0.94	0.95	0.99
SER	0.28	0.57	0.18	0.19	0.15	0.41	0.38	0.12
DW	2.33	1.25	2.13	2.03	1.89	1.92	1.89	1.93
DWH	0.98	0.58	0.43	0.48	0.75	0.01	0.01	0.83
ORT	0.59	0.02	0.07	0.37	0.11	0.05	0.51	0.92
WIT	0.57	18.68	22.05	2.04	7.72	11.41	7.67	10.45

Notes: IV estimates of the log-linear static demand equations. Dependent variable: log legal quantity sold. All regressors are in logs and real terms. All specifications include seasonal dummies, department fixed effects, and department-specific year effects. Wild cluster bootstrap standard errors, clustered by department, in parentheses. The Total column refers to aggregate legal sales in the observed segment. SER: standard error of the regression; DW: Durbin–Watson statistic; DWH and ORT: p-values of the Durbin–Wu–Hausman endogeneity test and the Hansen J overidentification test. WIT: the Lewis and Mertens (2025) weak-instrument statistic, increasing in instrument strength—values above the 5% critical value reject weak identification, values below indicate weak instruments. The critical value varies with the number of endogenous regressors and instruments and exceeds 20 throughout; the test fails to reject weak identification in all columns except DK3 (WIT = 22.05).

The IV evidence should therefore be interpreted as a diagnostic check rather than as the preferred estimator. It does not prove structural exogeneity of prices. Rather, it indicates that the available instruments do not provide a reliable alternative to the static conditional specification, while the endogeneity and overidentification diagnostics do not provide systematic evidence against conditioning on own prices. In Hendry's (1995) terminology, the relevant interpretation is one of weak exogeneity for the parameters of the conditional legal-demand equation, not structural exogeneity of prices.

Appendix B
COVID-period robustness

The sample includes the COVID-19 pandemic, which may have affected cigarette demand, retail activity, mobility, enforcement, distribution, and supply chains. To assess whether the main static demand results are driven by the pandemic period, Table B.1 re-estimates equation (13) including a COVID-period dummy equal to one between April and December 2020. The purpose is not to identify the causal effect of COVID-19, but to verify whether the estimated price responses are sensitive to controlling for the main pandemic-disruption period.

Table B.1

Static demand estimates with COVID control

Variable	DK1	DK2	DK3	BL1	BL2	BL3	BL4	Total
$p_{n,d,t}$	-2.11 (0.51)	-5.58 (0.96)	-2.37 (0.55)	-0.68 (0.39)	-1.14 (0.94)	-2.20 (0.44)	-1.97 (0.79)	-0.22 (0.24)
$P_{d,t}$	0.62 (0.36)	0.66 (0.55)	0.32 (0.23)	0.28 (0.36)	0.32 (0.10)	-0.43 (0.35)	0.48 (0.28)	0.23 (0.13)
p_t^I	0.49 (0.11)	1.39 (0.44)	0.54 (0.20)	0.47 (0.13)	0.29 (0.20)	1.67 (0.50)	-0.59 (0.31)	0.18 (0.13)
$p_{d,t}^{coffee}$	0.34 (0.18)	-0.48 (0.31)	0.16 (0.09)	-0.06 (0.12)	-0.07 (0.12)	-0.50 (0.17)	-0.21 (0.21)	0.05 (0.06)
$p_{d,t}^{beer}$	0.62 (0.23)	0.20 (0.43)	0.60 (0.24)	-0.60 (0.59)	-0.18 (0.18)	-1.61 (0.80)	0.48 (0.90)	-0.02 (0.17)
$p_{d,t}^{coca}$	-0.28 (0.09)	-0.45 (0.14)	-0.25 (0.04)	-0.03 (0.05)	-0.01 (0.06)	0.07 (0.25)	-0.29 (0.11)	-0.06 (0.04)
$p_{d,t}^{gum}$	-0.17 (0.19)	0.03 (0.16)	-0.01 (0.10)	-0.06 (0.07)	-0.02 (0.07)	-0.41 (0.41)	0.43 (0.22)	-0.08 (0.04)
y_t	1.63 (0.35)	2.72 (2.29)	0.24 (0.61)	0.59 (0.44)	0.53 (0.87)	2.17 (0.94)	-1.20 (0.90)	-0.26 (0.44)
COVID	0.09 (0.05)	0.01 (0.09)	0.20 (0.06)	-0.02 (0.03)	-0.02 (0.06)	-0.13 (0.07)	0.09 (0.05)	0.05 (0.04)
Observations	1161	1107	1161	1161	1161	1156	1161	1161
Adjusted R^2	0.98	0.96	0.99	0.99	0.98	0.95	0.96	0.99
SER	0.28	0.47	0.18	0.17	0.15	0.36	0.35	0.12
DW	2.33	1.93	2.10	2.10	1.89	1.91	2.15	1.93

Notes: OLS estimates of the log-linear static demand equations with a COVID dummy (equal to one from April to December 2020). Dependent variable: log legal quantity sold. Continuous regressors are in logs and, where applicable, real terms. All specifications include monthly seasonal dummies, department fixed effects, and department-specific year effects. Wild cluster bootstrap standard errors, clustered by department, in parentheses. The Total column refers to aggregate legal sales in the observed segment; its price variable is the sales-weighted average final price. SER: standard error of the regression; DW: Durbin–Watson statistic.

The results are very close to the baseline estimates. The own-price coefficients are negative for all brand-level equations; the magnitudes are similar to those in Table 1; and the aggregate legal demand elasticity remains small relative to the brand-level estimates. The central conclusion is therefore unchanged: aggregation conceals economically important substitution margins, and the aggregate legal elasticity is not a reliable guide to the vulnerability of the taxed legal base.

Appendix C

Variance decomposition of aggregate legal unit values

This appendix examines the extent to which changes in aggregate legal cigarette unit values reflect changes in within-brand prices or changes in brand composition. The exercise is motivated by the contrast between the aggregate and brand-level elasticities reported in Section 5.2. If aggregate unit values are driven mainly by substitution across brands, elasticities estimated from aggregated quantities and unit values may be poor guides to the response of the taxed base to excise changes.

Let the aggregate legal unit value in department d and month t be defined as:

$$UV_{d,t} = \frac{\sum_{n=1}^{N} q_{n,d,t} p_{n,d,t}}{\sum_{n=1}^{N} q_{n,d,t}} = \sum_{n=1}^{N} s_{n,d,t} p_{n,d,t},$$

where $q_{n,d,t}$ denotes legal sales of brand n in department d and period t, $p_{n,d,t}$ is the own real price of the brand, and

$$s_{n,d,t} = \frac{q_{n,d,t}}{\sum_{j=1}^{N} q_{j,d,t}},$$

is the share of brand n in total legal cigarette sales in department d and month t.

The change in the aggregate legal unit value between t-1 and t can then be written as:

$$\Delta UV_{d,t} = UV_{d,t} - UV_{d,t-1} = \sum_{n=1}^{N} \overline{s}_{n,d,t} \Delta p_{n,d,t} + \sum_{n=1}^{N} \overline{p}_{n,d,t} \Delta s_{n,d,t}, \tag{20}$$

where:

$$\overline{s}_{n,d,t} = \frac{s_{n,d,t} + s_{n,d,t-1}}{2}, \quad \overline{p}_{n,d,t} = \frac{p_{n,d,t} + p_{n,d,t-1}}{2},$$

and

$$\Delta s_{n,d,t} = s_{n,d,t} - s_{n,d,t-1}, \quad \Delta p_{n,d,t} = p_{n,d,t} + p_{n,d,t-1}.$$

Equation (20) is an exact midpoint decomposition. The first term captures the contribution of within-brand price changes, holding brand shares at midpoint values. The second term captures the contribution of brand-composition changes, holding prices at midpoint values. This distinction is central for the present application. If the second component dominates, then movements in aggregate legal unit values reflect primarily changes in brand shares rather than changes in prices within brands.

For notational convenience, define:

$$A_{d,t} = \sum_{n=1}^{N} \overline{s}_{n,d,t} \Delta p_{n,d,t}, \quad B_{d,t} = \sum_{n=1}^{N} \overline{p}_{n,d,t} \Delta s_{n,d,t},$$

so that

$$\Delta UV_{d,t} = A_{d,t} + B_{d,t}.$$

The variance of changes in aggregate legal unit values can therefore be decomposed as:

$$Var\left(\Delta UV_{d,t}\right) = Var\left(A_{d,t}\right) + Var\left(B_{d,t}\right) + 2Cov\left(A_{d,t}, B_{d,t}\right).$$

To evaluate the relative importance of each component, Table C.1 reports, for each department and for the national aggregate, the following normalized terms:

$$A = \frac{Var\left(A_{d,t}\right)}{Var\left(\Delta UV_{d,t}\right)}, \; B = \frac{Var\left(B_{d,t}\right)}{Var\left(\Delta UV_{d,t}\right)}, \; C = \frac{2Cov\left(A_{d,t}, B_{d,t}\right)}{Var\left(\Delta UV_{d,t}\right)},$$

where, by construction, $A + B + C = 1$.

The results in Table C.1 indicate that changes in aggregate legal cigarette unit values are driven more by changes in brand composition than by changes in within-brand prices. In every department, the variance share associated with the brand-composition component exceeds that associated with the within-brand price component. Moreover, the implied B/A ratio is often large. In Sucre, La Paz, Oruro, and Pando, the brand-composition component is more than three times as important as the within-brand price component; in Potosí and Beni it is close to three times as large; and even in Cochabamba and Tarija it remains clearly dominant. By contrast, once the data are aggregated nationally, the difference between the two components narrows substantially, with the aggregate ratio falling to little more than one. Thus, the disaggregated evidence indicates that brand

composition is often several times more important than within-brand price changes, whereas aggregation makes the two channels appear almost equally relevant.

Table C.1

Variance decomposition of changes in aggregate legal unit values

Department	Within-brand price (A)	Brand composition (B)	Covariance (C)
Sucre	0.198	0.693	0.108
La Paz	0.261	0.907	-0.168
Cochabamba	0.463	0.698	-0.161
Oruro	0.199	0.702	0.099
Potosí	0.279	0.733	-0.012
Tarija	0.379	0.582	0.039
Santa Cruz	0.597	0.741	-0.338
Beni	0.254	0.737	0.008
Pando	0.210	0.713	0.077
Total	0.628	0.787	-0.415

Notes: Variance decomposition of monthly changes in aggregate legal cigarette unit values (UV). "Within-brand price" (A) is the variance share from brand-level price changes holding brand shares at midpoint values; "Brand composition" (B) is the share from changes in brand shares holding prices at midpoint values; "Covariance" (C) is twice the covariance of the two components. By construction, $\mathbf{A} + \mathbf{B} + \mathbf{C} = 1$.

This has two related implications. First, the decomposition shows that an aggregate legal unit value is not primarily capturing a common price movement for comparable products. Instead, a substantial part of its variation reflects reallocation across brands. In that sense, the aggregate unit value is not simply a noisy proxy for a market price. It is a composite object that combines within-brand price changes and shifts in the composition of purchases across heterogeneous brands. The larger the degree of aggregation, the harder it becomes to recover the underlying source of variation. This is particularly clear in Table C.1: the departmental decompositions show that brand substitution is generally much more important than within-brand price changes, yet national aggregation compresses this difference and makes the two components appear almost equally relevant.

Second, once the data are aggregated and the relevant "price" is constructed as a weighted-average unit value, the data themselves lose much of their ability to reveal how important substitution across brands actually is. The decomposition is feasible here only because the underlying brand-level prices and quantities remain observable. If one started directly from aggregated unit values, the contribution of within-brand price changes and the contribution of brand substitution would be confounded in a single measure. In that sense, aggregation does not merely attenuate the role of substitution; it obscures the mechanism altogether.

This reading of the decomposition is also consistent with the economics of cigarette markets. It would be difficult to rationalize the observed variation in aggregate unit values as arising mainly from persistent price dispersion across locations for the same brand. Cigarettes are a tradable manufactured product. If the same brand were systematically sold at markedly different prices across departments, this would create incentives for arbitrage and resale, especially in an environment characterized by high informality. While transport costs, local distribution margins, market power, and enforcement differences may generate some degree of spatial price dispersion, it would be counterintuitive for this to account for most of the observed variation. The decomposition instead suggests that the dominant source of variation is reallocation across brands.

The implication for the main text is direct. In a differentiated-products market with substantial brand substitution, aggregate or cluster-level unit values are poor candidates for measuring the policy-relevant price variation induced by excise changes. They combine true price movements with changes in the mix of brands purchased and may also be affected by movement toward untaxed or under-taxed alternatives. As a result, elasticities estimated from aggregated quantities and unit values need not recover the response of the taxed base to excise reform. The results in Table C.1 suggest that such estimates would tend to understate the importance of substitution and thereby bias the analysis toward an overly optimistic view of excise increases.